第5章　非負行列因子分解を用いた大域信号除去

5.1　はじめに

前章では，ICA を用いた大域信号除去について述べた．そこで用いられたデータは，単一周波数 (223[Hz]) の環境電磁波の平均振幅であり，その信号の混合モデルには単純な線形和による混合を仮定していた．しかし，位相の無相関性を考慮した場合，線形混合モデルを当てはめるためには，厳密には，平均振幅ではなく平均エネルギーのデータを用いなければならない．加えて，現段階での ICA による解析には限界があると考えられる．現在までに提案されている ICA アルゴリズムの多くは，信号の独立性を評価する指標の関係から，0 を中心とする対称な確率分布に基づく源信号が生起していることを暗に仮定している．しかし，本研究で用いる ELF 帯環境電磁波データはエネルギーを表現したものであり，その源信号の平均値は 0 にならず，確率分布にも偏りがあると考えられる．そのような信号の独立性については，適切な定義が明らかでない．また，本研究では混合行列に非負性を仮定しているが，一般的な ICA にはその前提がない．このように，ICA が前提としている信号の性質と，実際の信号の物理的な性質との乖離が大きい可能性がある．そこで，より実際の物理現象に即した信号処理手法に置き換える必要がある．

扱う信号や混合過程に非負性を仮定している BSS 手法として，**非負行列因子分解 (Non-negative Matrix Factorization; NMF)** がある [13][14]．NMF とは，ある非負の行列を別の 2 つの非負行列の積へと近似的に因数分解する手法であり，主に非負値の多変量データを統計に基づいて解析する際に適用される．行列の因数分解は，数値的な線形代数において広く学ばれてきたが，これに非負の制約を加えることで，今まで解決が困難であった問題においても，適用できる可能性を見出せた．例えば，画像の重なり (映り込み) や音声スペクトルデータといった，全ての要素が非負値であり，また加算のみで表せるモデルにおいて，その解析を高い精度で行うことができると考えられる．

そこで本章では，第 3 章で説明した大域信号除去において，信号モデルの再定義を行い，大域信号の推定に NMF を用いた場合について述べる [34][35]．

5.2 観測信号の数理モデルの置き換え

ある源信号を $s_k(t)$ とすると，それがセンサ j で受信されたときの信号 $x_i(t)$ は，次のように書くことができる．

$$x_i(t) = \mathrm{AMP}\{g(s_k(t))\}\sin(\omega_0 t + \mathrm{PH}\{h(s_k(t))\} + \theta_0) \tag{5.1}$$

ここで，AMP，PH は信号の振幅と位相を表したもの，g，h は信号が受ける変調を関数として表したもの，ω_0，θ_0 は受信信号の角速度と初期位相である．

実際には無数に存在する放射源から電磁波が放射され，それぞれ異なった変調を伴って瞬時にセンサまで到達し，それらが重畳した状態の単一周波数成分 (223[Hz]) を我々が観測する．よって，式 (5.1) を次のように拡張して書きかえる．

$$x_i(t) = \sum_k A_{ik} s_k(t) \sin(\omega_0 t + \theta_k) \tag{5.2}$$

ここで，式 (5.1) における変調の関数 g，h は，源信号とセンサについてそれぞれブラックボックスであり，一般的には非線形関数と考えられるが，振幅の変調を線形な A_{ik}，位相の変調を初期位相を含めた固定値 θ_k として簡略化できる．また，単一周波数成分の加算であるため，角速度 ω_0 はセンサと源信号に依存しない定数 (223[Hz]) とみなせる．

式 (5.2) はセンサが得る信号であるが，データロガーに記録される情報は，6 秒間 (300 サンプル分) の $x_i(t)$ の絶対値平均である[1]．源信号の分離に際し，このデータロガーの値における各源信号がどのような状態かを，確率的な側面から考える．

源信号の生起過が互いに独立 (無相関) であると仮定すると，位相情報 θ_k は $0 \sim 2\pi$ の一様分布からの標本値とみなせることから，$x_i(t)$ の分布は平均 0，分散 $(A_{ik}s_k(t))^2/2$ の正弦分布を k に関して畳み込んだ分布となる．ここで，源信号は多数存在することから，中心極限定理より，$x_i(t)$ の分布は平均 0，分散 $\sum_k (A_{ik}s_k(t))^2/2$ 正規分布に置き換えることができる．そして，6 秒間 ($t = 300t' \sim 300t' + 299$) で分布が不変と仮定すると，データロガーの値は $x_i(t)$ の分布を $x_i = 0$ で折り返した分布 (半正規分布) の期待値とみなせる．半正規分布の期待値は元になった正規分布の標準偏差の定数倍であることから，データロガーの値 $x_i(t')$ は次のようになる．

$$x_i(t') = c\sqrt{E\left[\sum_k (A_{ik}s_k(t))^2/2\right]_{t=300t' \sim 300t'+299}} \tag{5.3}$$

ここで，c は比例係数，$E[\cdot]_{t=300t' \sim 300t'+299}$ は t' に対応する時刻 t の区間における期待値である．

[1] 50[Hz] で絶対値検波した後に，1/6[Hz] でダウンサンプリングしたものを記録する．

式 5.3 のモデルにおいて源信号の分離を行うのは困難である．そこで，まず，二乗することで平方根を外す．

$$x_i^2(t') = c^2 E\left[\sum_k (A_{ik}s_k(t))^2/2\right]_{t=300t' \sim 300t'+299} \tag{5.4}$$

次に，区間 $[300t', 300t'+299]$ の全ての t についての再現を諦め，その区間の代表値として分散 (平均エネルギー)$s_k^2(t')/2$ の分離を試みる．この場合，式 (5.4) は次のように書き変わる．

$$x_i^2(t') = c^2 \sum_k A_{ik}^2 s_k^2(t')/2 \tag{5.5}$$

最後に，変数を置き換え，定数を隠すと一般的な NMF の問題に置き換わる．

$$x_i^{'}(t') = \sum_k A_{ik}^{'} s_k^{'}(t') \tag{5.6}$$

5.3 非負行列因子分解 (NMF) の概要

5.3.1 問題の定式化

全ての要素が非負の $n \times T$ 行列 $\boldsymbol{X}$ が与えられた時，NMF アルゴリズムは

$$\boldsymbol{X} \approx \boldsymbol{AY} \quad \boldsymbol{A} \geq 0,\ \boldsymbol{Y} \geq 0 \tag{5.7}$$

となる $n \times r$ 行列 $\boldsymbol{A}$ と $r \times T$ 行列 $\boldsymbol{Y}$ の組を探す．ただし，$\boldsymbol{A}$，$\boldsymbol{Y}$ の要素は全て非負である．r は分解のランク数を表す任意の正の整数であるが，通常は不等式 $nT > nr + rT$ を満たす数を与える．このとき，出力される行列 $\boldsymbol{A}$，$\boldsymbol{Y}$ は与えられた行列 $\boldsymbol{X}$ よりも小さなサイズとなる．

5.3.2 推定アルゴリズム

現在，様々な NMF アルゴリズムが提案され，実装されている [36]．ここでは，従来研究で用いられていた，最も標準的なアルゴリズムである ISRA (Image Space Reconstruction Algorithm)[15][16] を説明する．

NMF は入力行列 $\boldsymbol{X}$ を出力行列の積 $\boldsymbol{AY}$ で表現することが目的であり，多くの NMF アルゴリズムは，$\boldsymbol{X}$ と $\boldsymbol{AY}$ との何らかの距離をコスト関数とし，その最小化を行う．ISRA では，$\boldsymbol{X}$ と $\boldsymbol{AY}$ の二乗ユークリッド距離 (二乗 Frobenius ノルム)

$$||\boldsymbol{X} - \boldsymbol{AY}||_F^2 = \sum_{ik}\{x_{ik} - [\boldsymbol{AY}]_{ik}\}^2 \tag{5.8}$$

をコスト関数として用いる．式 (5.8) 中の x_{ik} は行列 $\boldsymbol{X}$ の (i,k) 要素，$[\boldsymbol{AY}]_{ik}$ は行列の積 $\boldsymbol{AY}$ の (i,k) 要素を表している．

式 (5.8) を最小とするための更新式は，勾配法を用いると次の式 (5.9)(5.10) となる．

$$a_{ij} \leftarrow a_{ij} + \eta_{a_{ij}} \left\{ [\boldsymbol{X}\boldsymbol{Y}^T]_{ij} - [\boldsymbol{A}\boldsymbol{Y}\boldsymbol{Y}^T]_{ij} \right\} \tag{5.9}$$

$$y_{jk} \leftarrow y_{jk} + \eta_{s_{jk}} \left\{ [\boldsymbol{A}^T\boldsymbol{X}]_{jk} - [\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{Y}]_{jk} \right\} \tag{5.10}$$

ここで a，y はそれぞれ $\boldsymbol{A}$，$\boldsymbol{Y}$ の要素 (添え字は行列のインデックス) を，$\eta_{a_{ij}}$，$\eta_{y_{jk}}$ は勾配法における任意のステップ幅を表す.

勾配法において，ステップ幅が小さな正の数であれば，コスト関数は減少していく．ここで，式 (5.9)(5.10) には減算が含まれており，更新によって値が負になる可能性がある．しかし，NMF では値が負になってはならない．そこで，更新によって値が負にならないよう，ステップ幅を次のように定める．

$$\eta_{a_{ij}} = \frac{a_{ij}}{[\boldsymbol{A}\boldsymbol{Y}\boldsymbol{Y}^T]_{ij}} \tag{5.11}$$

$$\eta_{y_{jk}} = \frac{y_{jk}}{[\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{Y}]_{ij}} \tag{5.12}$$

この場合，更新式 (5.9)(5.10) はそれぞれ次の式 (5.13)(5.14) となる．

$$a_{ij} \leftarrow a_{ij} \frac{[\boldsymbol{X}\boldsymbol{Y}^T]_{ij}}{[\boldsymbol{A}\boldsymbol{Y}\boldsymbol{Y}^T]_{ij}} \tag{5.13}$$

$$y_{jk} \leftarrow y_{jk} \frac{[\boldsymbol{A}^T\boldsymbol{X}]_{jk}}{[\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{Y}]_{jk}} \tag{5.14}$$

これは乗算による更新である．よって，$\boldsymbol{X}$，$\boldsymbol{A}$，$\boldsymbol{Y}$ の要素が全て非負であれば，更新によって値が負になることはない．また，更新の際に $\boldsymbol{A}$ の各要素を次の式 (5.15) で標準化する．

$$a_{ij} = \frac{a_{ij}}{\sum_j a_{ij}} \tag{5.15}$$

ISRA では，式 (5.13),(5.14),(5.15) による更新を十分な回数 (通常は 500 回) 繰り返す．また，0 による除算を避けるため，値の更新において 10^{-16} 未満の要素を 10^{-16} に切り上げる．

NMF は解を一意に算出することができない．NMF は通常，$\boldsymbol{A}$，$\boldsymbol{Y}$ それぞれを非負の乱数で初期化し，Mote Carlo シミュレーションを適用して得られた解の中から最適なものを選び，最終的な解とすることが多い．しかし本研究では，解の一意性と解析結果の良好さから，特異値分解を用いた初期化 [14] を行う．

アルゴリズムの収束性について，簡潔に説明する．ある関数 $f(x)$ の $x = x^t$ における補助関数 $g(x, x^t)$ は，

$$g(x, x^t) \leq f(x) \tag{5.16}$$

$$g(x, x) = f(x) \tag{5.17}$$

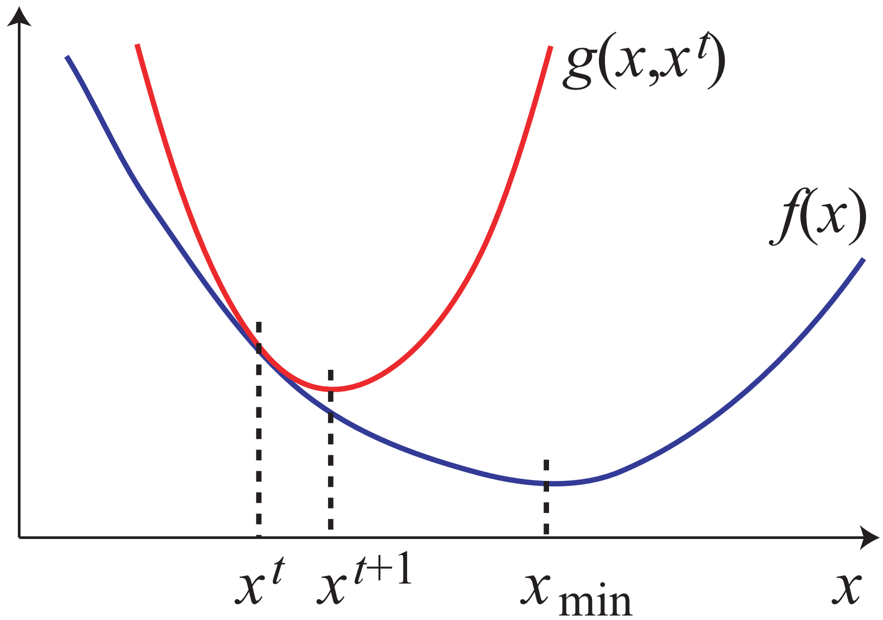

図 5.1: 補助関数による最適化

を満たすとする (図 5.1)．このとき，f と g は

$$f(x^{t+1}) \leq g(x^{t+1}, x) \leq g(x^t, x^t) = f(x) \tag{5.18}$$

という関係を満たすため，

$$x^{t+1} \leftarrow \arg\min_t g(x, x^t) \tag{5.19}$$

なる更新において f は非増加である．

5.3.3　時系列信号への適用

式 (7.6) において，入力行列 $\boldsymbol{X}$ から t 列目の要素をベクトル $\boldsymbol{x}$ として取り出し，出力行列 $\boldsymbol{Y}$ から t 列目の要素をベクトル $\boldsymbol{y}$ として取り出すと，$\boldsymbol{x}$ と $\boldsymbol{y}$ は

$$\boldsymbol{x} \approx \boldsymbol{A}\boldsymbol{y} \tag{5.20}$$

という関係を満たす．よって，$n \times T$ 行列 $\boldsymbol{X}$ は T 列の n 次元列ベクトル，$r \times T$ 行列 $\boldsymbol{Y}$ は T 列の r 次元列ベクトル，$n \times r$ 行列 $\boldsymbol{A}$ は $\boldsymbol{Y}$ を $\boldsymbol{X}$ へと線形変換する作用素と捉えることができる．このことから，NMF を時系列信号に適用すると，観測信号 $\boldsymbol{x}(t) = [x_1(t), ..., x_n(t)]^T$ を非負の混合行列 $\boldsymbol{A}$ と非負の信号成分 $\boldsymbol{y}(t) = [y_1(t), ..., y_r(t)]^T$ の積へと，近似的に因数分解する．つまり

$$\boldsymbol{x}(t) \approx \boldsymbol{A}\boldsymbol{y}(t) \quad t = 1, 2, ..., T \tag{5.21}$$

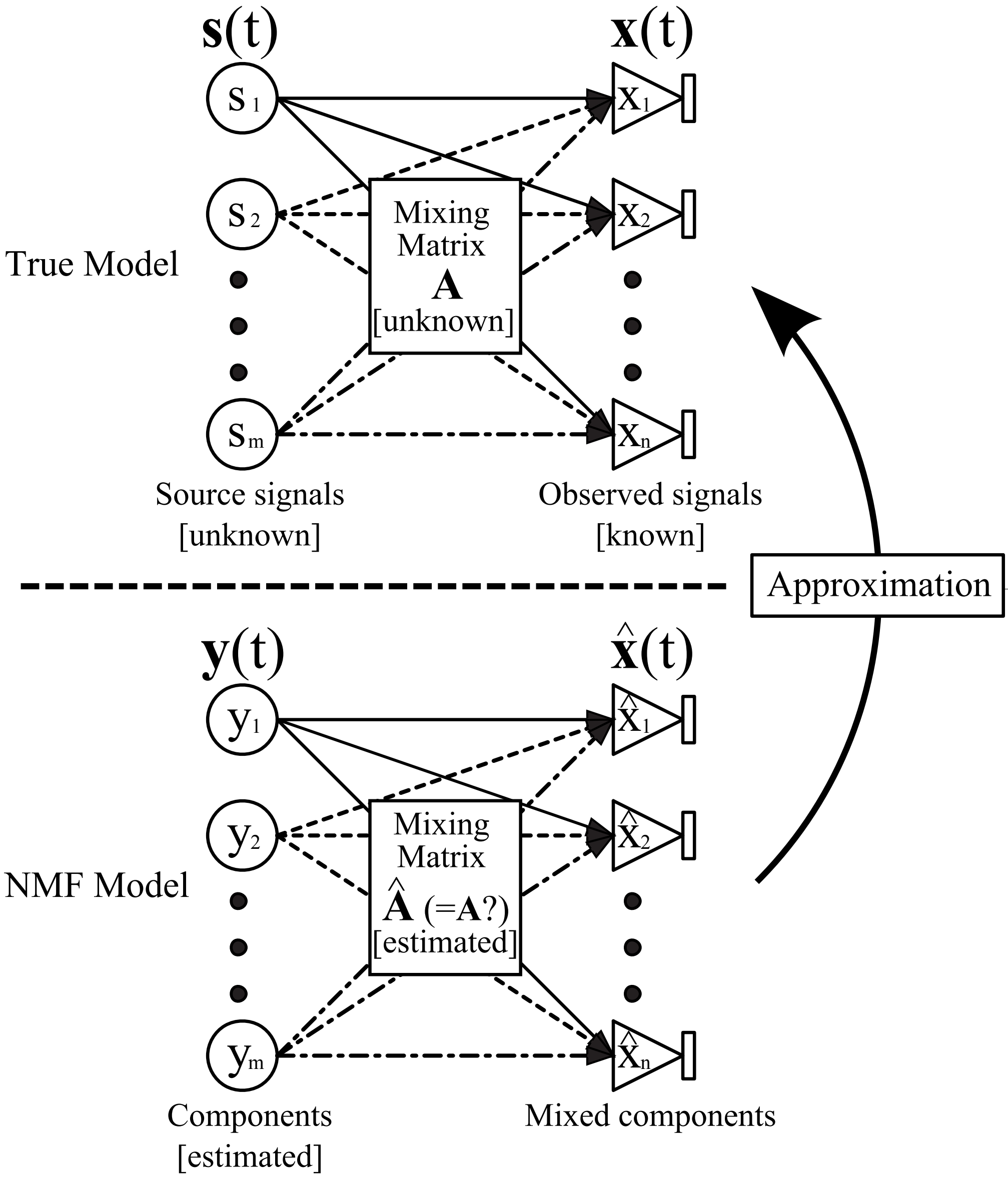

図 5.2: 時系列信号での NMF モデル

となる．ただし，$\boldsymbol{A} \geq 0$，$\boldsymbol{y}(t) \geq 0$ である．

式 (5.21) は，式 (4.3) との比較から，源信号とその混合モデルを推定することと同じ意味を持つ (図 5.2)．つまり，NMF を非負の時系列信号に適用することで，非負の源信号とその混合モデルを推定することができる．

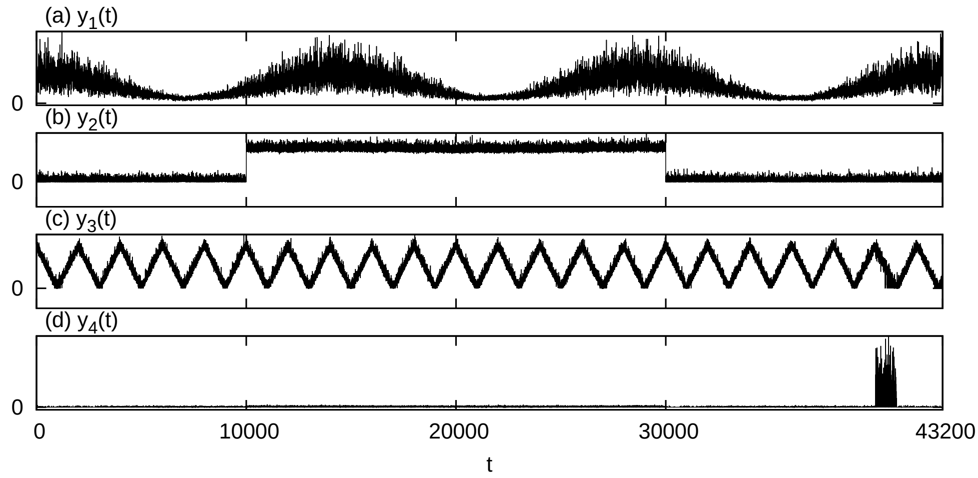

図 5.3: ISRA によって推定された源信号

表 5.1: 推定された局所信号の SN 比および GIC

データ	SNR[dB]	GIC
観測信号	-5.073	1.002
推定された局所信号		
理想	—	0.054
平均的な観測信号	6.471	0.120
NG-FICA (全ての観測信号を解析)	7.911	0.278
NG-FICA (選択した観測信号を解析)	13.400	0.061
ISRA (全ての観測信号を解析)	17.533	0.047

5.4 シミュレーション実験

NMF を用いた大域信号除去が可能であることを確認するため．ELF 帯環境電磁波データを模した擬似信号に対し，大域信号除去の実験を行う．第 4.6 節のものと同じ擬似信号に対して行い，ICA を用いた場合と結果を比較する．

ISRA によって推定された源信号を図 5.3 に示す．分解のランク数 r は 4 とした．$y_1(t)$ から $y_4(t)$ は，推定の順序が入れ替わっているものの，真の源信号によく似ている．そして，ICA の場合とは異なり，波形の上下が逆転しているものはない．$y_1(t)$ は大域信号である $s_1(t)$ を推定したものと断定できるため，これを除去の対象とする．

ISRA によって推定された局所信号を図 5.4 に示す．観測信号に見られた 3 サイクルの波が見られなくなっており，理想的な局所信号 (4.4) とよく似ていることから，大域信号が除去されていることが予想される．今回の実験では，真の局所信号を非負値のみで構成しているが，ISRA によって推定された局所信号には，小さな負値が確認できる．これは，ISRA が誤差行列 $\boldsymbol{E} \equiv \boldsymbol{X} - \boldsymbol{AY}$ に非負の制約を設けていないためである．$\boldsymbol{X}$，$\boldsymbol{A}$，$\boldsymbol{Y}$ が非負値のみで構成

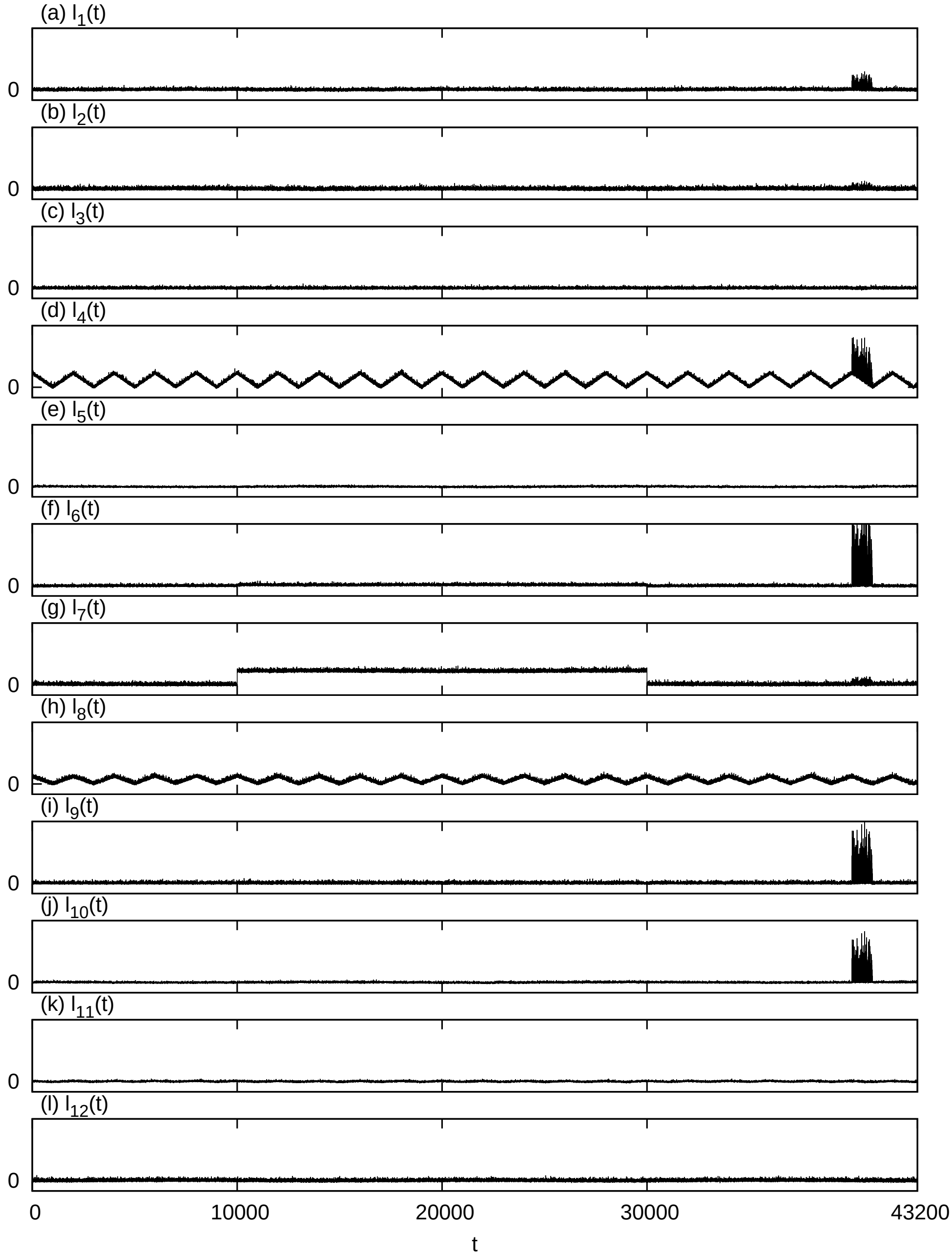

図 5.4: ISRA によって推定された局所信号

されている場合でも，減算を伴う $\boldsymbol{E}$ は負値を取りうる．センサごとに異なる雑音が加わった状態においては，完璧な因子分解は不可能であり，大域信号除去など成分の減算を伴う処理を行った場合に負値が表れたとしても，その値が 0 に近い場合は許容する必要がある．

推定された局所信号の精度を，SN 比および GIC に基づいて評価する (表 5.1)．表より，ISRA によって推定された局所信号の SN 比は，NG-FICA によって推定された局所信号よりも改善しており，GIC も小さくなっている．このことから，ISRA は NG-FICA よりも ELF 帯環境電磁波の解析に適している可能性が高いといえる．

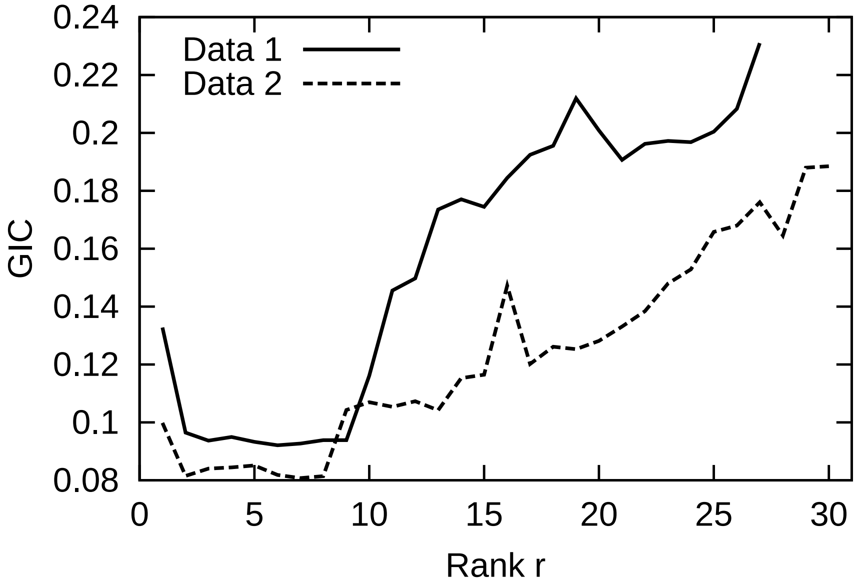

図 5.5: ISRA を用いた場合の分解のランク数 r と推定局所信号の GIC の関係

5.5 ELF 帯環境電磁界観測データへの適用

5.5.1 分解のランク数 r の設定

NMF の適用に際し，分解のランク数を適切に設定することは重要である．通常，分解のランク数は想定される源信号の数に相当するが，ELF 帯環境電磁波の信号源は無数に存在することから，分解のランク数を非常に大きく設定することが必要と考えられる．しかし，実際には複数の観測点にまたがって観測される電磁波信号は非常に少なく，単一の観測点でのみ観測される電磁波信号は，振幅の小さなものであれば，誤差の中に含めることが可能である．よって，分解のランク数は小さな値でもよい．

後述の実験においては，分解のランク数を r をさまざま変えて実験を行った．その結果を図 5.5 に示す．図の横軸は分解のランク数 r，縦軸は r に対応する GIC を表しており，実線が実験 1 の場合，破線が実験 2 の場合である．$r = 6$ の場合に GIC が最も小さくなる傾向がみられたため，後述の実験に関しては $r = 6$ の場合の結果を載せる．

5.5.2 実験 1

2005 年 3 月 17 日に，長崎県雲仙の観測点において異常信号を確認した．その 3 日後である 3 月 20 日に福岡県西方沖地震 (M7.0) が発生していることから，確認された異常信号が地震と関連している可能性がある．そこで，3 月 16 日から 18 日にかけての観測信号に対し，大域信号除去を行う．

表 5.2: 大域信号除去結果の評価 (2005年3月16日から18日，二乗値利用)

データ	*GIC*
観測信号	0.2343
推定された局所信号	
平均的な観測信号	0.1155
NG-FICA (全ての観測信号を解析)	0.2337
NG-FICA (選択した観測信号を解析)	0.1007
ISRA (全ての観測信号を解析)	0.0910

観測信号を図5.6に示す．図の (a) から (d) は，それぞれ (a) 秋田県男鹿，(b) 岐阜県萩原，(c) 大阪府茨木，(d) 長崎県雲仙において観測された信号の二乗値を表している．図の横軸は時系列 (日時) を，縦軸はELF帯環境電磁界のエネルギー [$\mathrm{pT}^2/\mathrm{Hz}$] を表している．(d) 長崎県雲仙の観測信号には，3月16日の1時頃から12時頃の振幅が非常に大きくなっていることが確認できる．また，大域信号の特徴である夜は大きく昼は小さいといった日変化が，全ての観測信号において共通して確認できる．

図5.7は平均的な観測信号である．解析する観測信号の選択や，推定された源信号からの大域信号の特定は，この信号との相互相関係数に基づいて行う．

全ての観測信号からNG-FICAによって推定された大域信号と局所信号を図5.8と図5.9に，選択した観測信号からNG-FICAによって推定された大域信号と局所信号を図5.10と図5.11に，ISRAによって推定された大域信号と局所信号を図5.22と図5.13に示す．図の横軸は時系列 (日時) を，縦軸は信号の振幅を表している．

図5.8に示す大域信号は，日変化らしき特徴を持っているが，負の方向に大きな値がいくつか見られる．大域信号が適切に推定できていない可能性が高い．実際に，図5.9に示す推定局所信号は，依然として共通な成分を含んでいるように見える．なお，大域信号はICAによって推定されたにもかかわらず平均が0でないが，これは，復元行列 $\boldsymbol{W}$ を推定する際にはあらかじめ平均を減じた観測信号を用い，復元信号 $\boldsymbol{y}(t)$ を計算する際には平均を減じていない観測信号を用いたためである．

選択した観測信号からNG-FICAによって推定された大域信号 (図5.10) は，全ての値が負となった．これは，ICAが持つスケールの不定性と，非負の制約が設けられていないことが原因であり，波形の推定自体には成功しているように見える．実際に，図5.11に示す推定局所信号は，共通な成分である日変化がほとんど見られず，大域信号除去に成功しているように見える．

ISRAによって推定された大域信号 (図5.22) は，スケール以外に関して，選択した観測信号からNG-FICAによって推定された大域信号 (図5.10) とほぼ同等の結果となった．また，

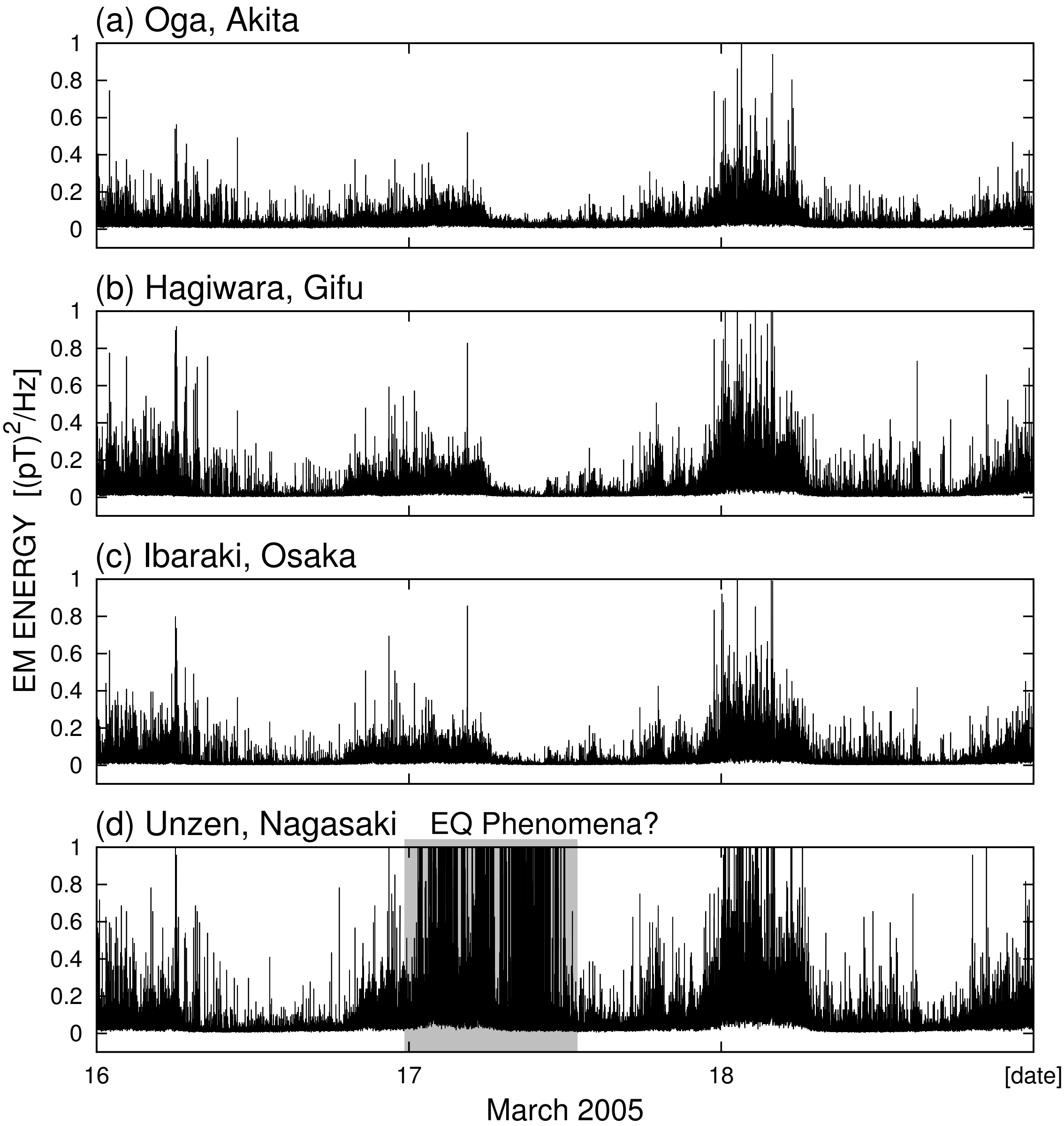

図 5.6: ELF 帯環境電磁波の観測信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

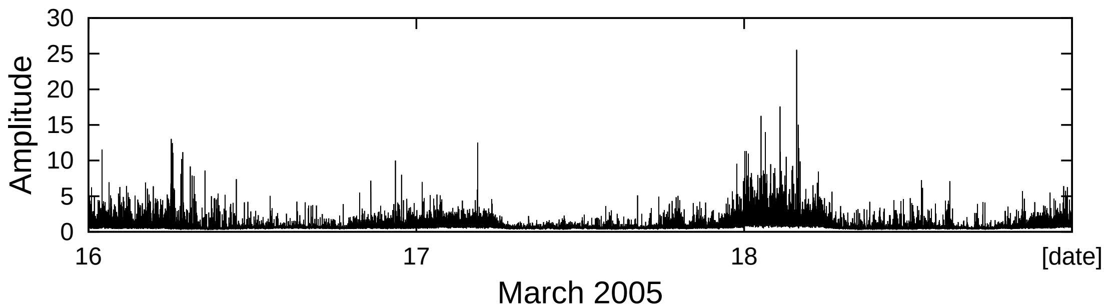

図 5.7: 平均的な観測信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

推定局所信号 (図 5.13) も共通な成分がほとんど見られず，大域信号が適切に除去されているように見える．

それぞれの手法によって推定された局所信号を，GIC に基づいて評価する．表 5.2 は，観測信号およびそれぞれの局所信号から求めた GIC である．GIC が小さいほど大域信号が適切に除去されていることを意味する．全ての観測信号から NG-FICA を用いて推定した局所信号は，GIC がほとんど小さくなっていない．大域信号除去に失敗しているといえる．選択した観測信号から NG-FICA を用いて推定した局所信号は，GIC が小さくなっている．ISRA を用いて推定した局所信号は，GIC がより小さくなっている．このことから，ELF 帯環境電磁波データの解析には，NG-FICA より ISRA の方が適している可能性が高い．

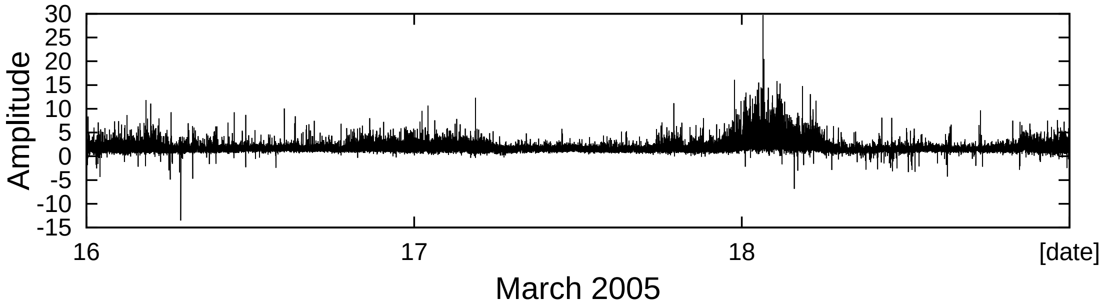

図 5.8: 全ての観測信号から NG-FICA によって推定された大域信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

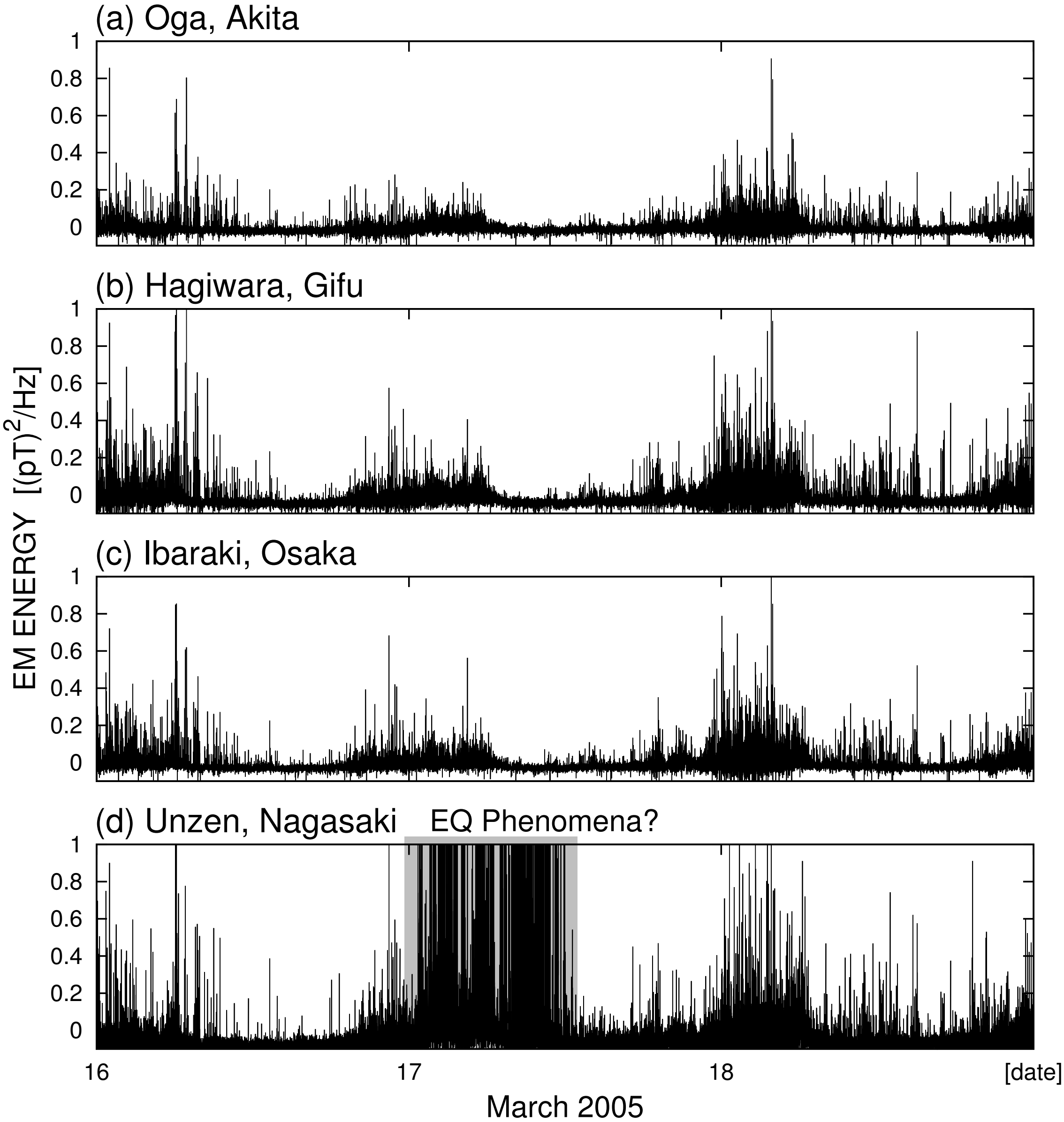

図 5.9: 全ての観測信号から NG-FICA によって推定された局所信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

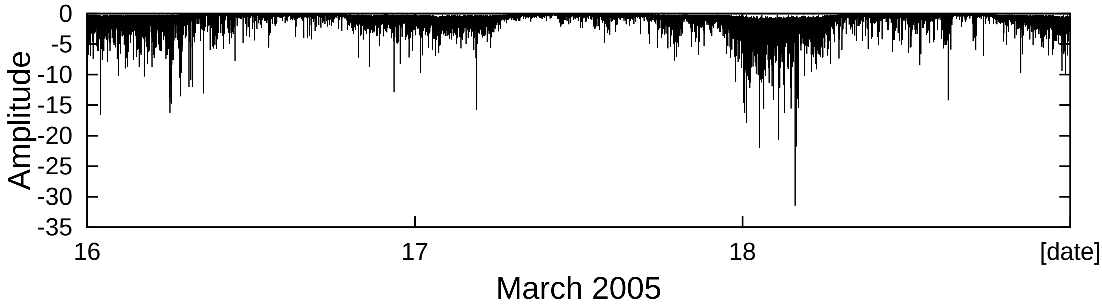

図 5.10: 選択した観測信号から NG-FICA によって推定された大域信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

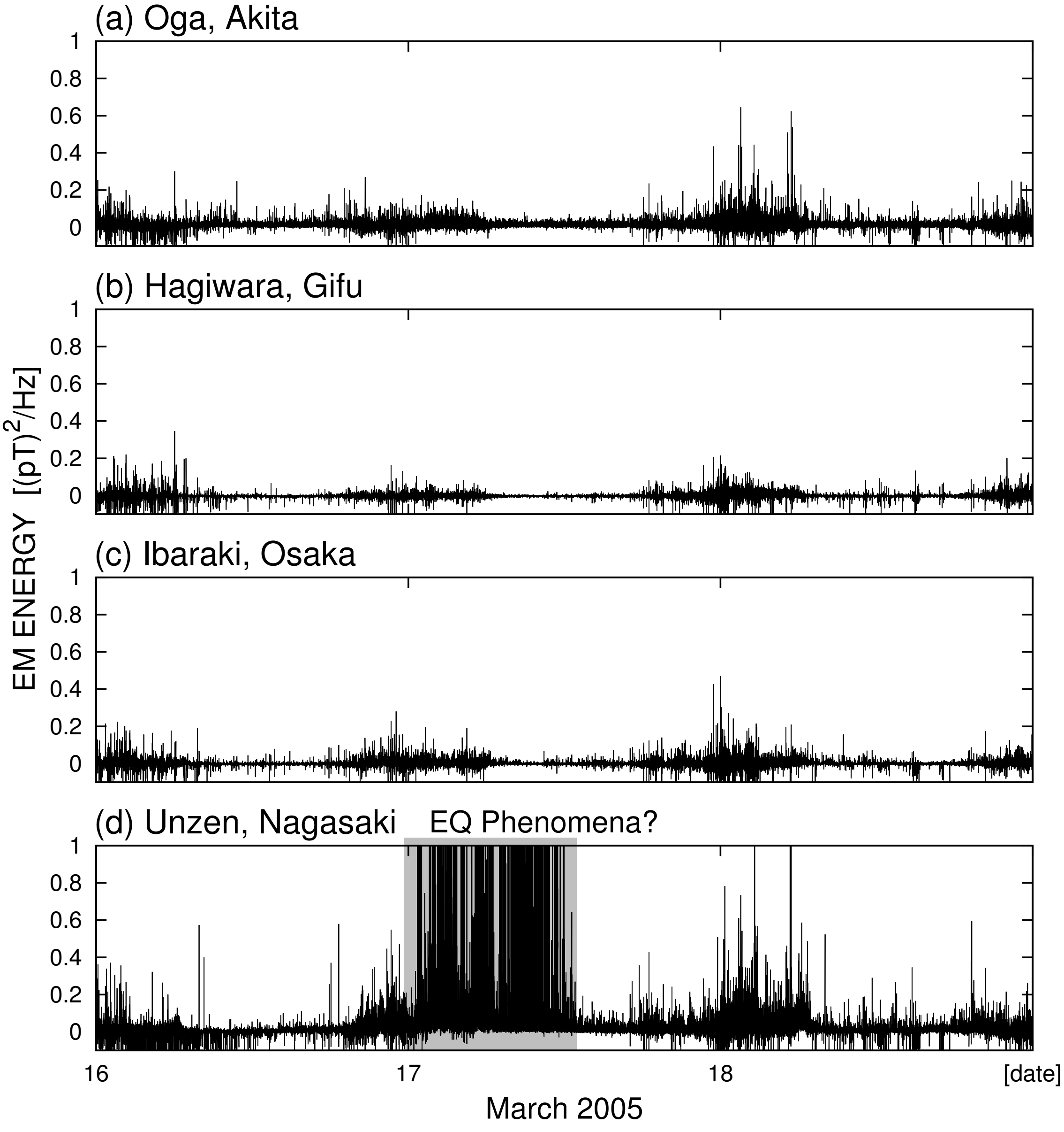

図 5.11: 選択した観測信号から NG-FICA によって推定された局所信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

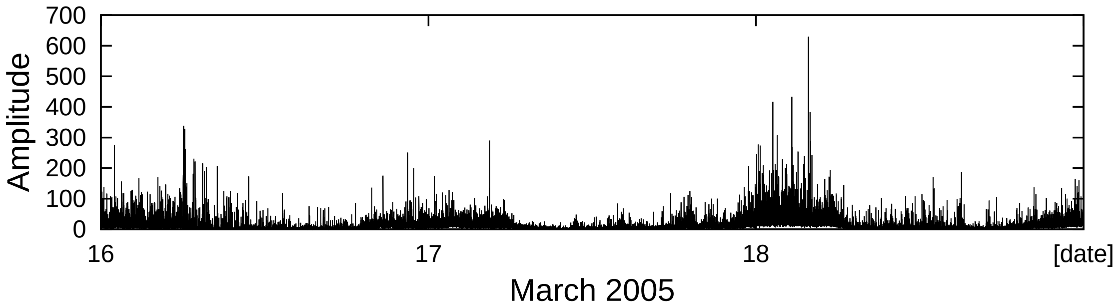

図 5.12: ISRA によって推定された大域信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

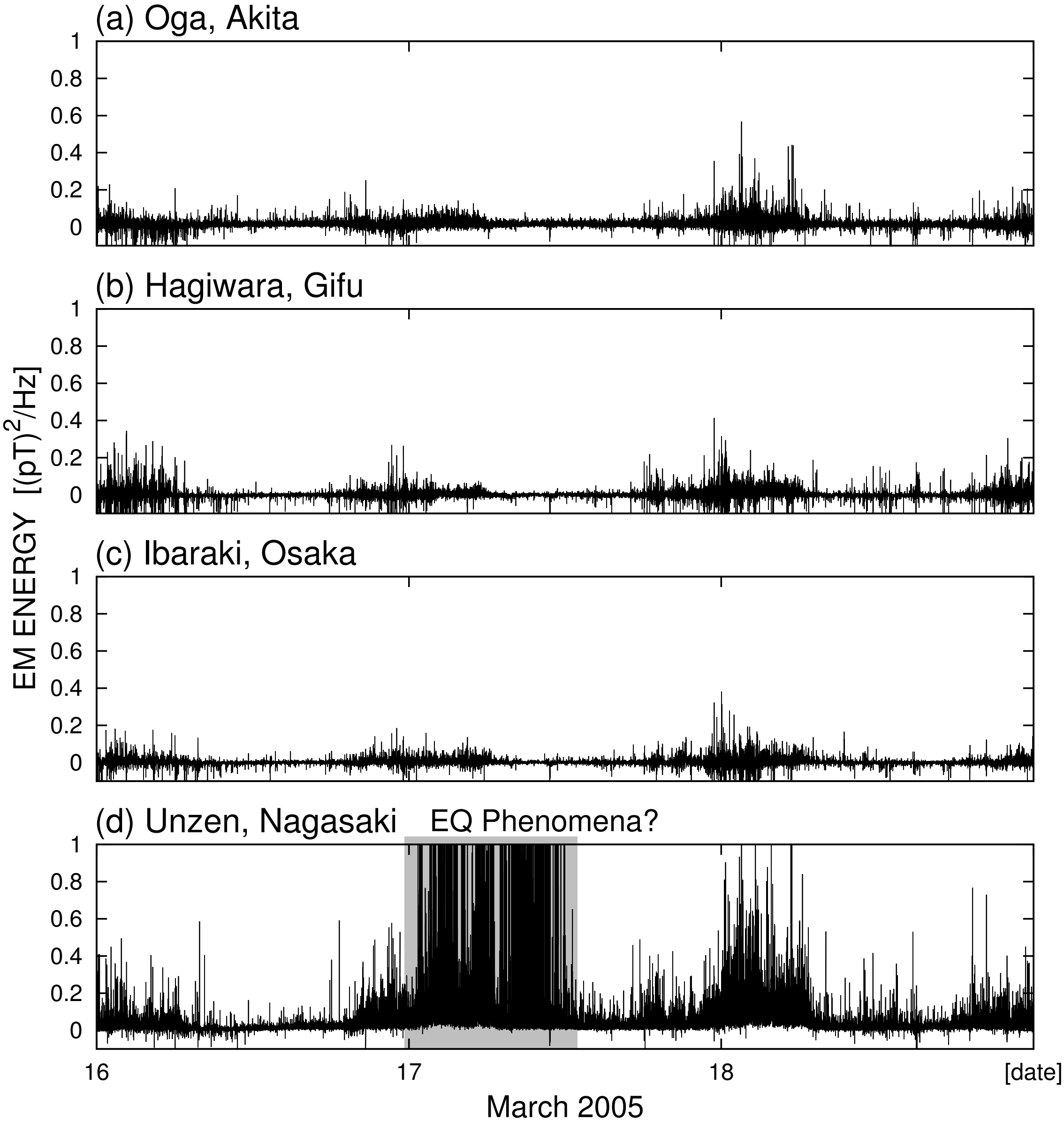

図 5.13: ISRA によって推定された局所信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

表 5.3: 大域信号除去結果の評価 (2001年1月4日から6日，二乗値利用)

データ	*GIC*
観測信号	0.1893
推定された局所信号	
平均的な観測信号	0.0958
NG-FICA (全ての観測信号を解析)	0.2157
NG-FICA (選択した観測信号を解析)	0.0781
ISRA (全ての観測信号を解析)	0.0809

5.5.3 実験2

2001年1月4日から6日にかけて，岐阜県南濃の観測点において異常信号が観測された．同月6日11時48分に岐阜県東濃においてM4.8の規模の地震が発生しており，岐阜県南濃で観測された異常信号との関連が考えられる．そこで，2001年1月4日〜6日の観測信号に対し，大域信号除去を適用する．

2001年1月4日〜6日に観測された信号のうち，一部の観測点のものを図5.14に示す．図の(a)から(d)は，それぞれ(a)秋田県男鹿，(b)岐阜県坂内，(c)岐阜県南濃，(d)長崎県雲仙に設置した観測点に対応している．岐阜県南濃の観測信号(図5.14(c))は，他の観測信号と比較しても，異常信号を確認することは難しい．他の観測点と外形が似ていることから，大域信号に埋もれてしまっているといえる．

図5.15は平均的な観測信号である．解析する観測信号の選択や，推定された源信号からの大域信号の特定は，この信号との相互相関係数に基づいて行う．

全ての観測信号からNG-FICAによって推定された大域信号と局所信号を図5.16と図5.17に，選択した観測信号からNG-FICAによって推定された大域信号と局所信号を図5.18と図5.19に，ISRAによって推定された大域信号と局所信号を図5.23と図5.21に示す．横軸は時系列(日時)を，縦軸は信号の振幅を表している．

図5.16に示す大域信号は，日変化らしき特徴を持っている．しかし，図5.17に示す推定局所信号は，依然として共通な成分を含んでいるように見える．大域信号が適切に推定できていない可能性が高い．

選択した観測信号からNG-FICAによって推定された大域信号(図5.18)は，実験1と同様に，全ての値が負となったが，波形の推定自体には成功しているように見える．図5.19に示す推定局所信号には，共通な成分である日変化がほとんど見られない．また，観測信号では確認が難しかった(c)の岐阜県南濃の異常信号が，局所信号では明瞭になっている．実験2においても大域信号除去にも成功しているように見える．

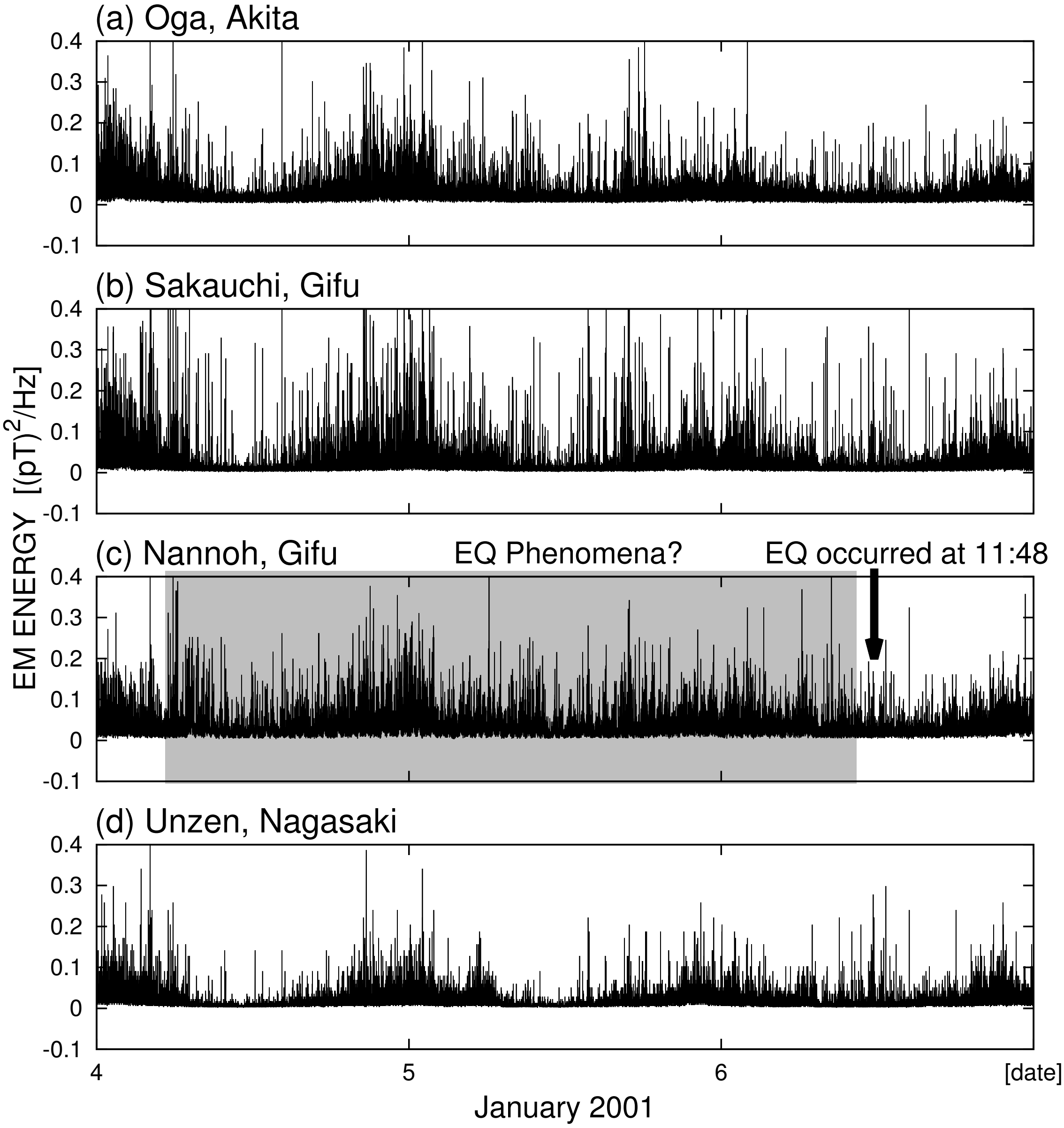

図 5.14: ELF 帯環境電磁波の観測信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

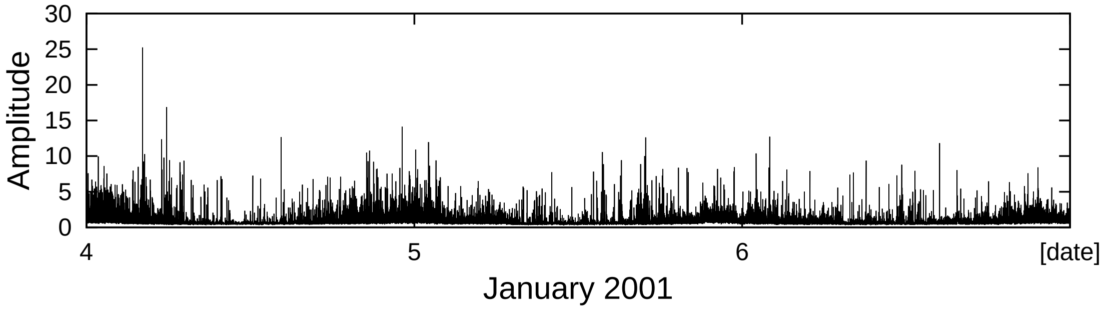

図 5.15: 平均的な観測信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

ISRA によって推定された大域信号と局所信号 (図 5.23 と図 5.21) も，実験 1 と同様に，大域信号のスケール以外に関して，選択した観測信号から NG-FICA によって推定された信号 (図 5.18 と図 5.19) とほぼ同等の結果となった．

それぞれの手法によって推定された局所信号から求めた GIC を，表 5.3 に示す．GIC が小さいほど大域信号が適切に除去されていることを意味する．実験 2 では，全ての観測信号から NG-FICA を用いて推定した局所信号は，GIC が大きくなった．大域信号除去に失敗していることが明らかである．選択した観測信号から NG-FICA を用いて推定した局所信号は，GIC が小さくなっており，ISRA を用いて推定した局所信号から求めた GIC は，それよりも若干大きくなっていた．よって，実験 2 においては，選択した観測信号から NG-FICA を用いて推定した場合が，最も高い精度で大域信号の推定と除去ができており，ISRA を用いて推定した場合はそれと同程度の結果が得られたといえる．

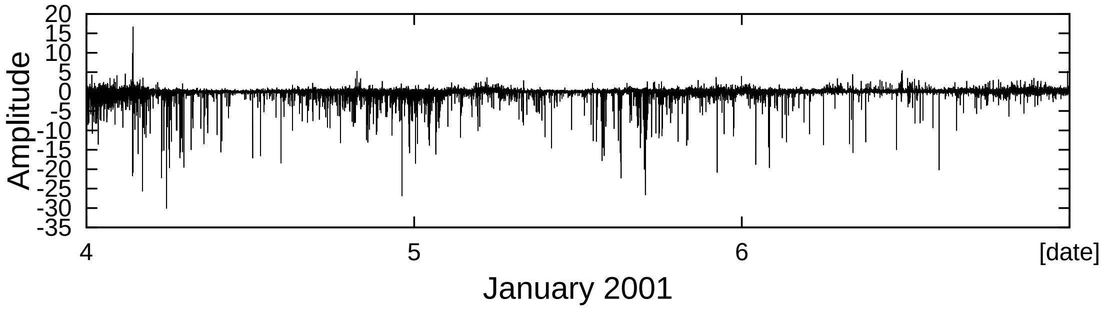

図 5.16: 全ての観測信号から NG-FICA によって推定された大域信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

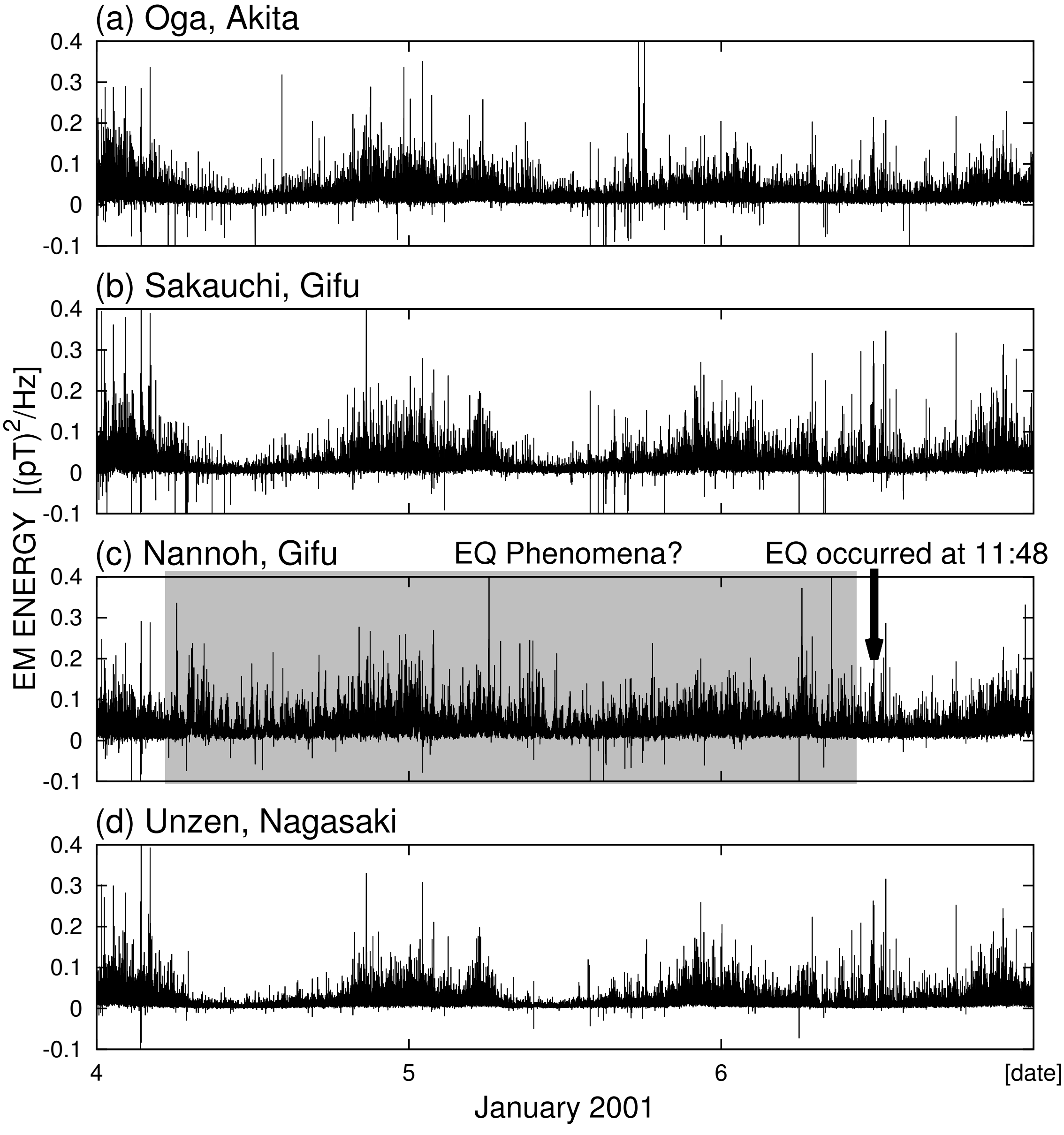

図 5.17: 全ての観測信号から NG-FICA によって推定された局所信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

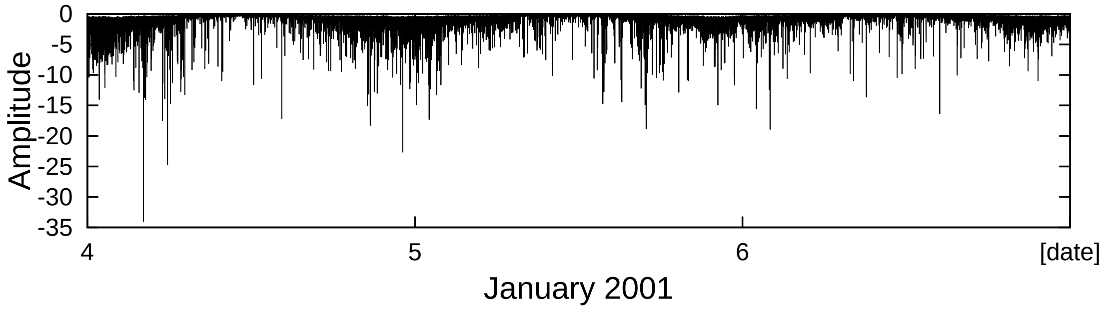

図 5.18: 選択した観測信号から NG-FICA によって推定された大域信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

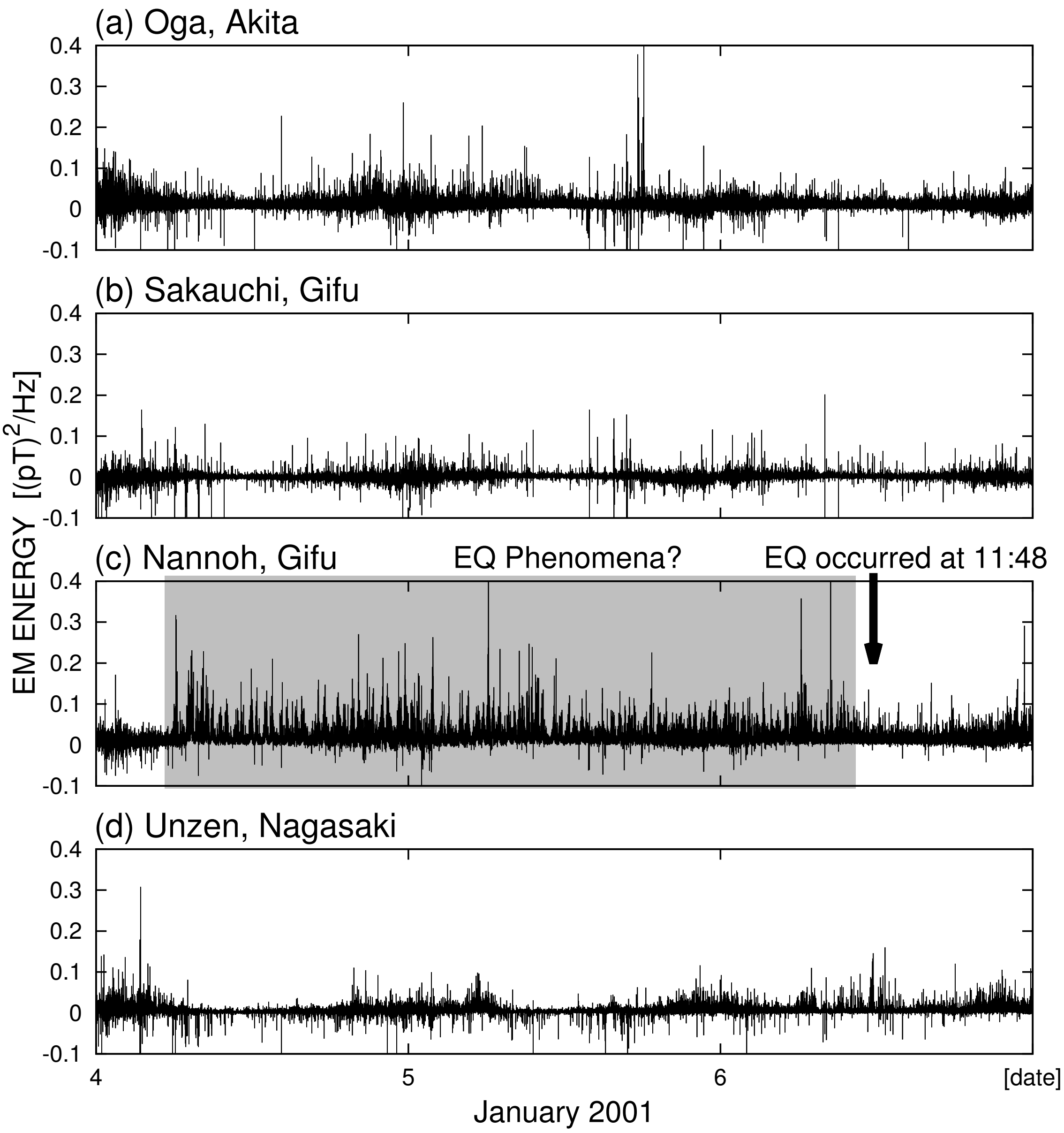

図 5.19: 選択した観測信号から NG-FICA によって推定された局所信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

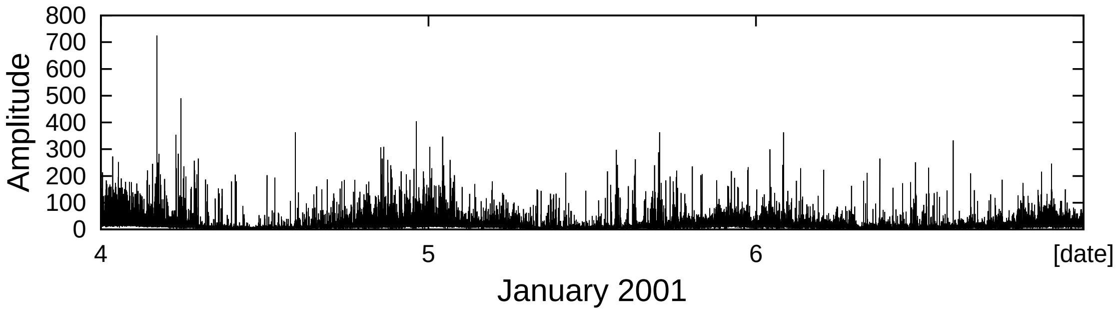

図 5.20: ISRA によって推定された大域信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

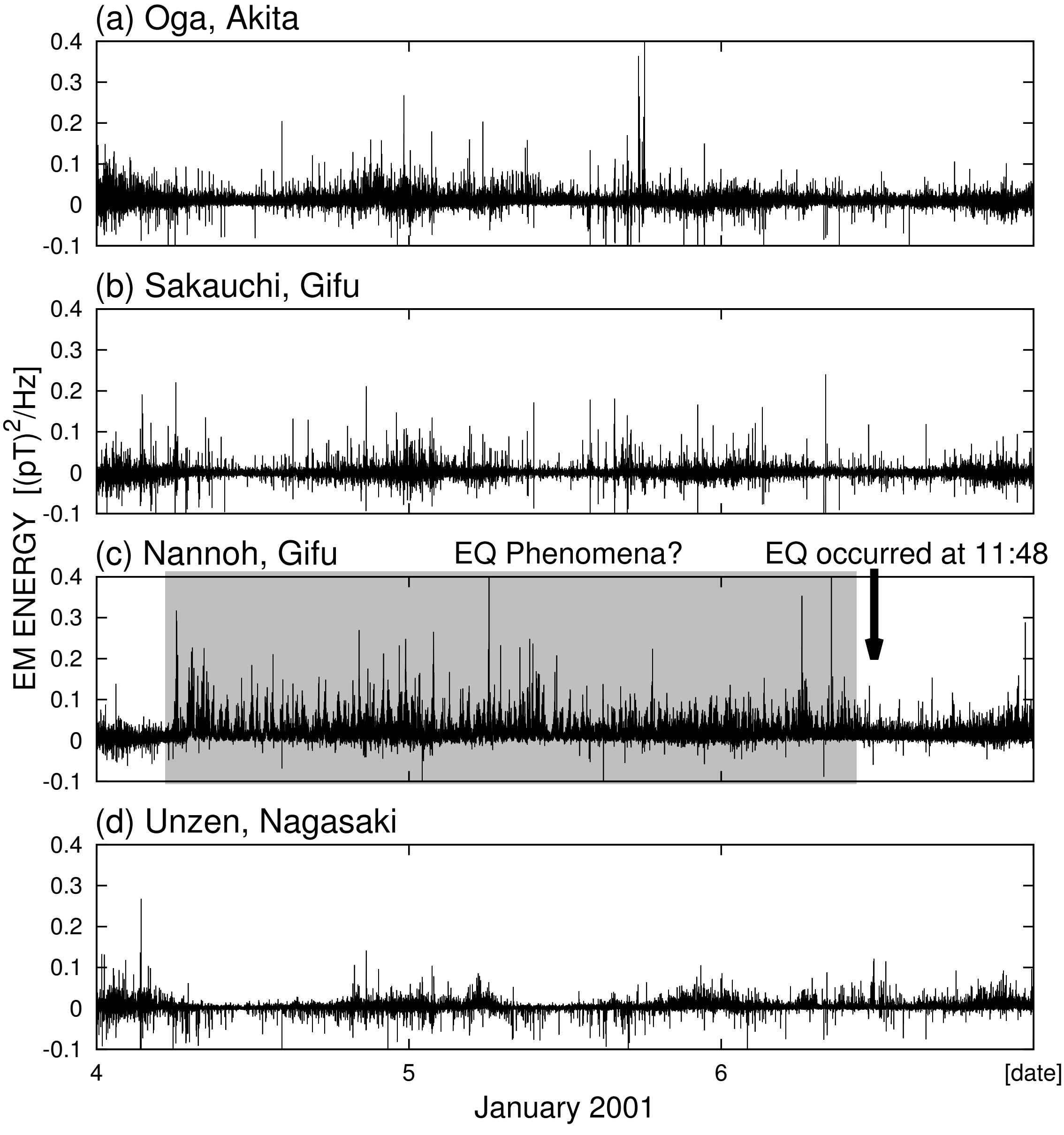

図 5.21: ISRA によって推定された局所信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

5.5.4　考察

実験1と実験2より，ISRAを用いて推定した大域信号と局所信号は，選択した観測信号からNG-FICAを用いて推定したものと同程度の推定精度が期待できることが分かった．ここで，ISRAは解析する観測信号を限定するという工夫を取っておらず，大域信号の感度ベクトル$\boldsymbol{b}$も混合行列$\boldsymbol{A}$の一部として大域信号と同時に推定している．その上で，推定精度がほぼ同程度であることから，ELF帯環境電磁波データの解析手法としてISRAはNG-FICAの妥当な代替であるといえる．また，大域信号が高い精度で推定されているのであれば，他の源信号も高い精度で推定されている可能性も考えられる．

実験1と実験2においてISRAによって推定された，大域信号以外の源信号を，ぞれぞれ図5.22と図5.23に示す．図の横軸は時系列(日付)を，縦軸は信号の振幅を表し，各図の(b)から(f)は，それぞれ$y_2(t)$からy_6を表す．また，ISRAによって推定された混合行列$\boldsymbol{A}$を，図5.24と図5.25に示す．左奥へ延びる軸は推定された源信号のインデックスを，右奥へ延びる軸は観測点のインデックスを，縦軸は混合行列$\boldsymbol{A}$の値(対数スケール)を表している．最も手前の列(赤色)はほとんどの観測点でほぼ等しい値であり，それ以外の列(水色)は一様でないことから，どちらも1番目の成分が大域信号であることが分かる．

図5.22(e)は，長崎県雲仙の観測信号(図5.6(d))と似ている．また，混合行列$\boldsymbol{A}$(図5.24)に関しても，5番目の成分におけるNo.21の観測点(長崎県雲仙)に対応する値が突出している．よって，この成分は長崎県雲仙において観測された異常信号の源信号を推定したものと言える．

図5.23には，岐阜県南濃において観測された異常信号と似ているものはなかった．異常信号が源信号として推定されなかった原因は，複数の観測点にまたがって観測されておらず，異常信号の振幅も大きくないためと考えられる．

5.6　まとめ

本章では，大域信号の推定にNMFを用いた場合について述べた．まず，ICAが想定している信号モデルと実際のELF帯環境電磁波観測信号との相異について述べ，信号モデルの非負の制約から，源信号推定にはICAよりNMFが適合していることを述べた．次に，NMFの概要とそのアルゴリズムであるISRAについて説明した．さらに，シミュレーション実験と実際のELF帯環境電磁波を用いた実験を行い，NMFがICAと同等以上の精度で大域信号を除去できることを示した．同時に，全ての観測信号を解析対象とできるNMFの利点から，地震前兆信号を直接抽出できる可能性について述べた．

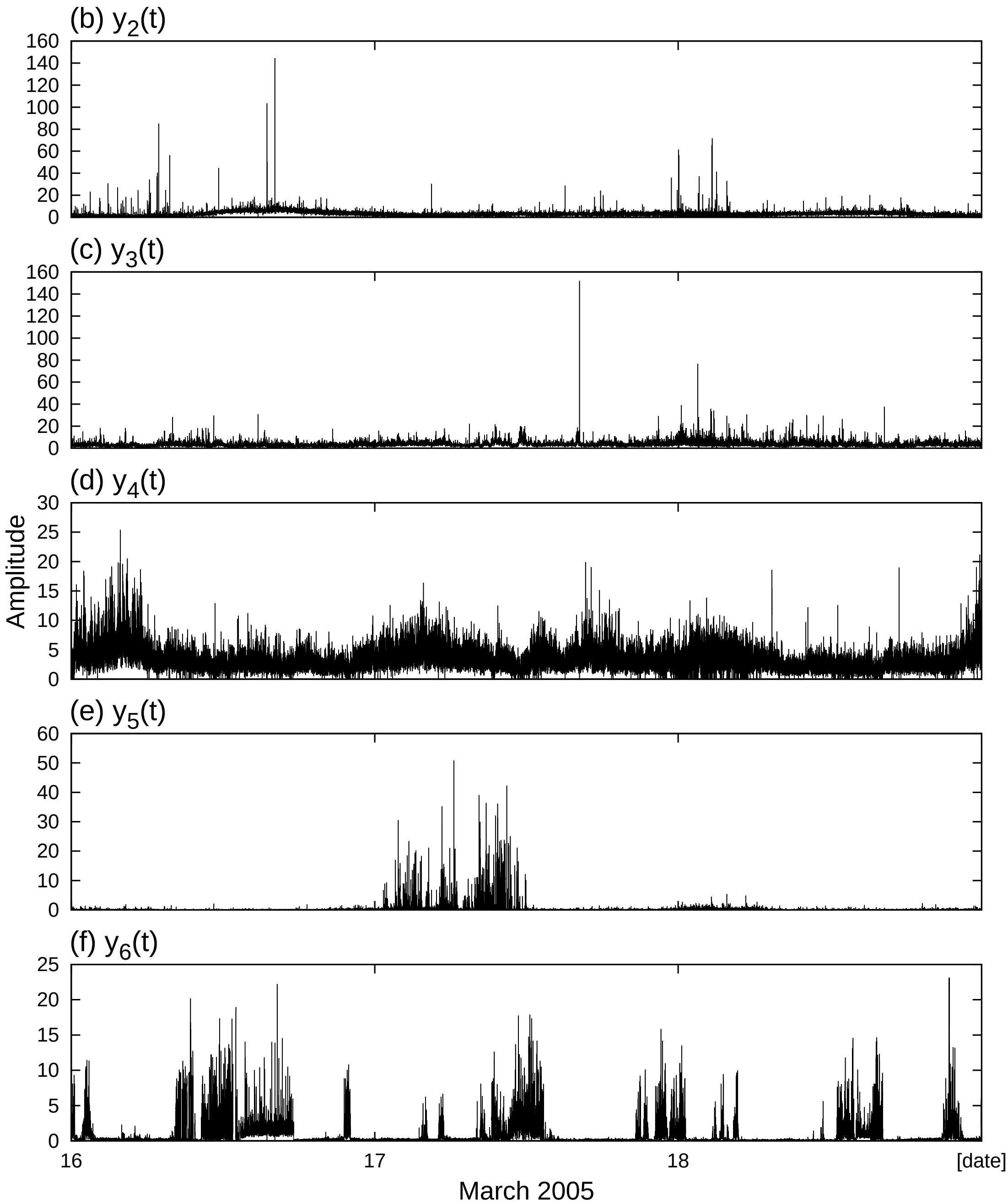

図 5.22: ISRA によって推定された，大域信号以外の源信号 (2005 年 3 月 16 日から 18 日，二乗値利用)

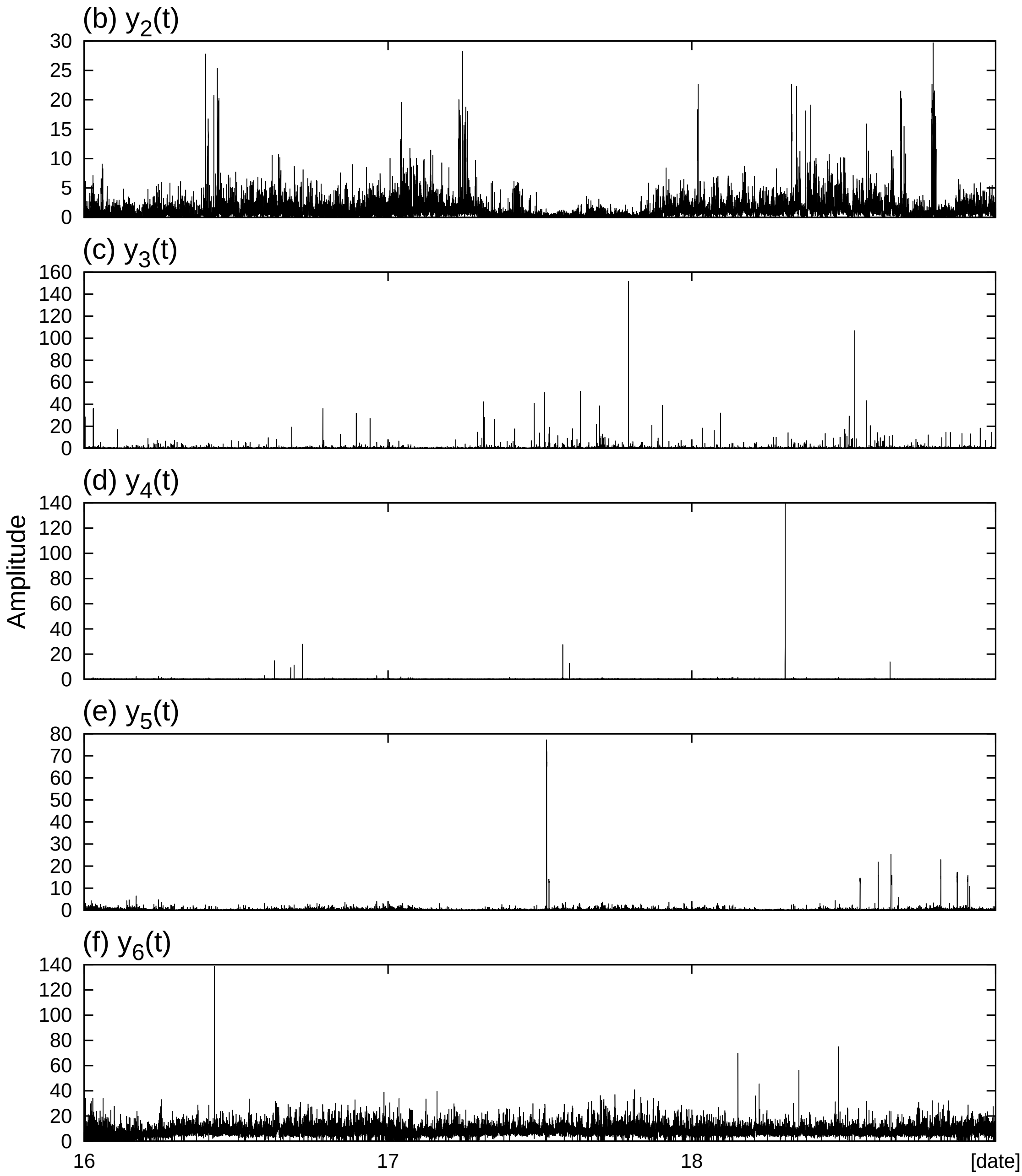

図 5.23: ISRA によって推定された，大域信号以外の源信号 (2001 年 1 月 4 日から 6 日，二乗値利用)

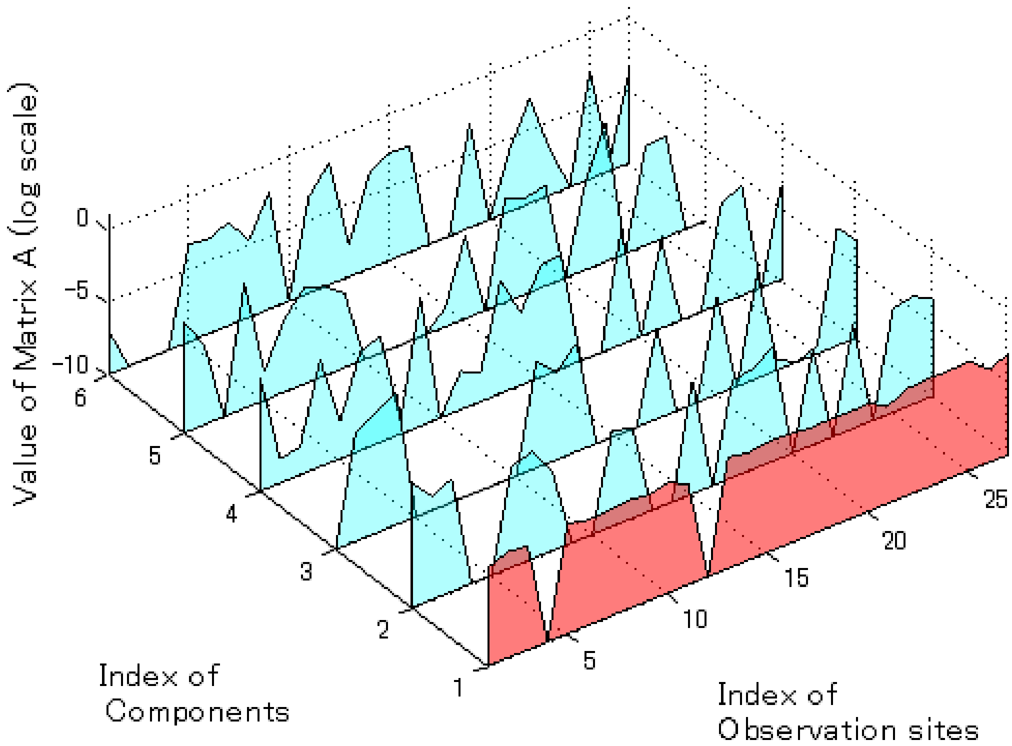

図 5.24: ISRA によって推定された混合行列 (2005 年 3 月 16 日から 18 日，二乗値利用)

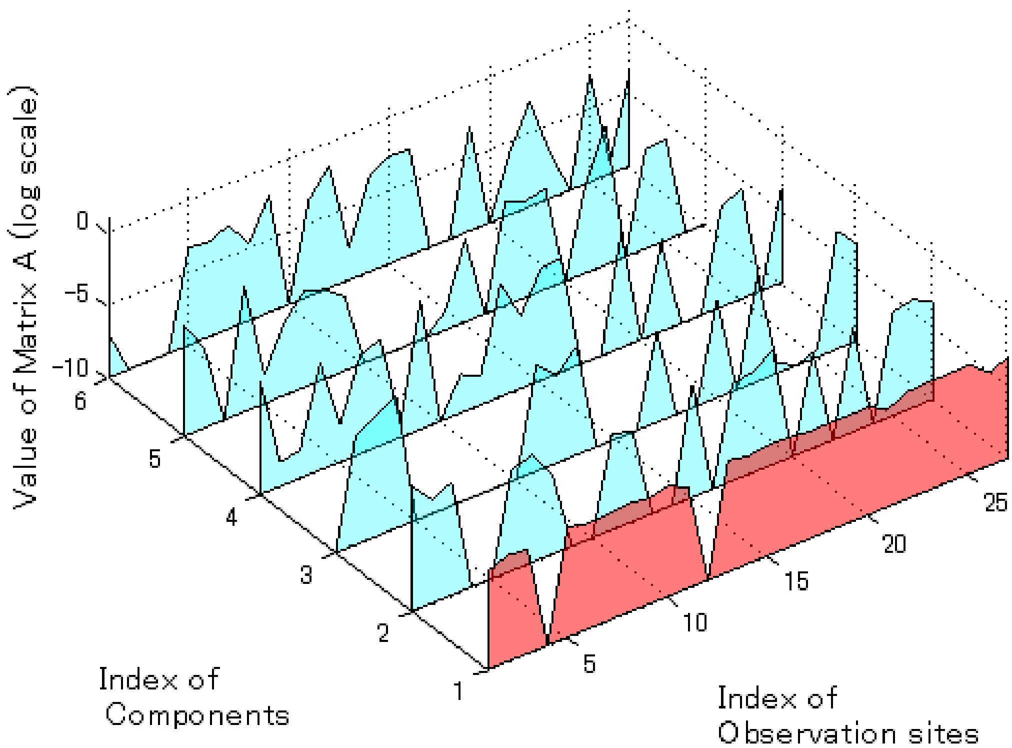

図 5.25: ISRA によって推定された混合行列 (2001 年 1 月 4 日から 6 日，二乗値利用)

第6章　擬似L1ノルム最小化に基づく非負行列因子分解を用いた源信号推定

6.1　はじめに

第5章では，大域信号の推定と除去に，NMFの基本的なアルゴリズムであるISRA[15][16]を用いていた．しかし，ISRAは解析結果が外れ値に影響されやすく，外れ値が多く含まれる信号の解析に失敗する場合があることが判明した[37]．ELF帯環境電磁波の観測信号には，インパルス状の信号やスパイク状の信号が記録される場合が多く，それらが原因で源信号や局所信号の推定精度が下がることは，回避すべき問題である．外れ値に対する頑強性を向上するための工夫が求められる．

逐次推定手法においては，一般に，コスト関数が低次である方が外れ値に頑強になる．ISRAはコスト関数がL2ノルム(二乗距離)に基づいており，これよりも低次のコスト関数に基づいた，新しいNMFアルゴリズムを考える．本研究では，擬似的なL1ノルムをコスト関数へと導入し，それに基づいた更新則を導出する[38]．本研究に類似した研究として，L1ノルムをコスト関数としたテンソル積展開が提案されているが[39]，背景信号(大域信号)の推定に特化したものである．本研究では，大域信号に限らず，源信号の推定そのものを目指す．

6.2　擬似L1ノルムを最小化する非負行列因子分解手法

次のような行列因子分解を仮定する．

$$\boldsymbol{X} \approx \boldsymbol{A}\boldsymbol{S} \tag{6.1}$$

ここで，$\boldsymbol{X}$は$n \times T$行列，$\boldsymbol{A}$は$n \times r$行列，$\boldsymbol{S}$は$r \times T$行列である．列のインデックスを時系列とみなしたとき，式(6.1)を次のように捉えなおすことができる．

$$\boldsymbol{x}(t) \approx \boldsymbol{A}\boldsymbol{s}(t) \quad (t = 1, 2, ..., T) \tag{6.2}$$

ここで，$\boldsymbol{X}$の各列が観測信号$\boldsymbol{x}(t)(t = 1, 2, ..., T)$からなるとき，$\boldsymbol{S}$の各列は源信号$\boldsymbol{s}(t)$に，$\boldsymbol{A}$はその混合係数行列に対応付けることができる．つまり，行列分解手法を用いて時系列多

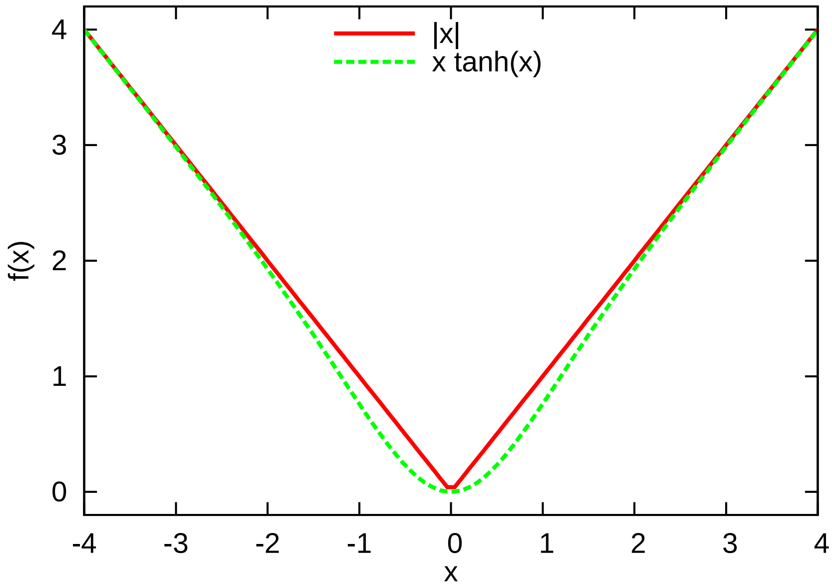

図 6.1: 擬似絶対値距離関数

変量データの成分推定ができる (図 5.2)．観測信号を $\boldsymbol{X}$ として入力した場合，源信号は $\boldsymbol{S}$，混合行列は $\boldsymbol{A}$ として出力される．

勾配法を用いて式 (6.1) を解く場合，コスト関数となる行列間距離関数には微分不可能な点を持つ絶対値演算を利用できない．L1 ノルムを用いる場合，原点で微分不可能な絶対値演算が含まれるため，勾配法を用いることができない．そこで，誤差行列 $\boldsymbol{E} = \boldsymbol{X} - \boldsymbol{AS}$ の評価関数 D を，絶対値演算を近似した $x\tanh(x)$(図 6.1) を用いて

$$D(\boldsymbol{E}) \equiv \sum_{i,k} E_{ik}\tanh(E_{ik}) \tag{6.3}$$

のように定義し，これを最小化するような $\boldsymbol{A}$ と $\boldsymbol{S}$ を求める．

式 (6.3) に対する A_{ij} の勾配は，

$$\begin{aligned}\frac{\partial}{\partial A_{ij}} D(\boldsymbol{E}) &= \sum_{l,k}\left\{\left(\frac{\partial}{\partial A_{ij}} E_{lk}\right)\cdot\left(\frac{\partial}{\partial E_{lk}} E_{lk}\tanh(E_{lk})\right)\right\} \\ &= \sum_{l,k}\left\{\left(\frac{\partial}{\partial A_{ij}}\sum_{m}(X_{lm} - A_{lm}S_{mk})\right)\cdot\left(\tanh(E_{lk}) + \frac{E_{lk}}{\cosh^2(E_{lk})}\right)\right\} \qquad (6.4)\end{aligned}$$

ここで，偏微分により $l = i$，$m = j$ の項のみ残ることから

$$\begin{aligned} &= \sum_{l,k}\left\{-S_{jk}\left(\tanh(E_{ik}) + \frac{E_{ik}}{\cosh^2(E_{ik})}\right)\right\} \\ &= -\sum_{k} S_{jk}\left\{\tanh(E_{ik}) + \frac{E_{ik}}{\cosh^2(E_{ik})}\right\} \qquad (6.5)\end{aligned}$$

となる．すなわち，式 (6.3) をコスト関数としたときの $\boldsymbol{A}$ の最急降下法における更新式は，次のようになる．

$$A_{ij} \leftarrow A_{ij} + \eta_A \sum_k S_{jk} \left\{ \tanh(E_{ik}) + \frac{E_{ik}}{\cosh^2(E_{ik})} \right\} \tag{6.6}$$

同様に，式 (6.3) をコスト関数としたときの $\boldsymbol{S}$ の最急降下法における更新式は，次のようになる．

$$S_{jk} \leftarrow S_{jk} + \eta_S \sum_i A_{ij} \left\{ \tanh(E_{ik}) + \frac{E_{ik}}{\cosh^2(E_{ik})} \right\} \tag{6.7}$$

ここで，η_A，η_S はステップ幅を表す．

式 (6.6) と (6.7) は単純な加算的更新則であり，更新によっては負の値を取り得る．そこで，これらの更新則をさらに拡張することで，非負の更新則を得る．まず，更新則の一部は次のように書き換えることができる．

$$\begin{aligned} \tanh(E_{ik}) + \frac{E_{ik}}{\cosh^2(E_{ik})} &= \frac{\sinh(E_{ik})\cosh(E_{ik})}{\cosh^2(E_{ik})} + \frac{E_{ik}}{\cosh^2(E_{ik})} \\ &= \frac{\exp(2E_{ik}) - \exp(-2E_{ik})}{4\cosh^2(E_{ik})} + \frac{4(X_{ik} - [\boldsymbol{AS}]_{ik})}{4\cosh^2(E_{ik})} \\ &= \frac{\exp(2E_{ik}) + 4X_{ik}}{4\cosh^2(E_{ik})} - \frac{\exp(-2E_{ik}) + 4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})} \end{aligned} \tag{6.8}$$

ここで，X_{ik}，$[\boldsymbol{AS}]_{ik}$，$\exp(\cdot)$ はすべて必ず非負値であるため，第一項は必ず正の値もしくは 0 となり，第二項は必ず負の値もしくは 0 となる．

式 (6.8) を代入すると，$\boldsymbol{A}$ の更新則は次のように書き換えられる．

$$A_{ij} \leftarrow A_{ij} + \eta_A \sum_k S_{jk} \frac{\exp(2E_{ik}) + 4X_{ik}}{4\cosh^2(E_{ik})} - \eta_A \sum_k S_{jk} \frac{\exp(-2E_{ik}) + 4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})} \tag{6.9}$$

このとき，η_A を拡張することで，非負の項を取り除くことができる．η_A を，$0 < \alpha_A < 1$ なる係数 α_A を用いて $\boldsymbol{A}$ の各要素に対してそれぞれ

$$\eta_{A_{ij}} \equiv \alpha_A A_{ij} \left\{ \sum_k S_{jk} \frac{\exp(-2E_{ik}) + 4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})} \right\}^{-1} \tag{6.10}$$

のように与えたとき，式 (6.9) の右辺は次のようになる．

$$\begin{aligned} & A_{ij} + \eta_{A_{ij}} \sum_k S_{jk} \frac{\exp(2E_{ik}) + 4X_{ik}}{4\cosh^2(E_{ik})} - \eta_{A_{ij}} \sum_k S_{jk} \frac{\exp(-2E_{ik}) + 4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})} \\ = \;& A_{ij} + \frac{\alpha_A A_{ij} \sum_k S_{jk} \frac{\exp(2E_{ik})+4X_{ik}}{4\cosh^2(E_{ik})}}{\sum_k S_{jk} \frac{\exp(-2E_{ik})+4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})}} - \alpha_A A_{ij} \end{aligned} \tag{6.11}$$

よって，$\boldsymbol{A}$ の更新則は次のようになる．

$$A_{ij} \leftarrow (1 - \alpha_A) A_{ij} + \frac{\alpha_A A_{ij} \sum_k S_{jk} \frac{\exp(2E_{ik})+4X_{ik}}{4\cosh^2(E_{ik})}}{\sum_k S_{jk} \frac{\exp(-2E_{ik})+4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})}} \tag{6.12}$$

このとき，X_{ik}，A_{ij}，S_{jk} が全て非負，かつ，$0 < \alpha_A < 1$ であれば，更新を行っても A_{ij} は非負値を維持する．同様に，$\boldsymbol{S}$ の非負の更新則は次のようになる．

$$S_{jk} \leftarrow (1-\alpha_S)S_{jk} + \frac{\alpha_S S_{jk} \sum_i A_{ij} \frac{\exp(2E_{ik})+4X_{ik}}{4\cosh^2(E_{ik})}}{\sum_i A_{ij} \frac{\exp(-2E_{ik})+4[\boldsymbol{AS}]_{ik}}{4\cosh^2(E_{ik})}} \tag{6.13}$$

最適な α_A と α_S は解析対象の性質によって異なると考えられる．ELF 帯環境電磁波に関しても未調査であるが，本研究では $\alpha_A = 1/2$，$\alpha_S = 1/2$ と置いて推定を行うこととする．また，一般的な NMF と同様に，Quasi-L1 NMF も解を一意に算出することができず，解の収束先が初期値の影響を大きく受ける．そこで，本研究では ISRA と同様に特異値分解を用いた初期化 [14] を行う．

6.3　源信号の推定精度の検証

計算機シミュレーションにより，源信号の推定精度の検証を行う．まず，図 6.2 に示す 4 つの源信号 $\boldsymbol{s}(t) = [s_1(t), ..., s_4(t)]^T$ $(t = 1, 2, ...14400)$ を作成した．図の縦軸は信号の振幅を，横軸は時系列に相当する t を表す．これらの源信号は ELF 帯環境電磁波信号を模した擬似信号であり，$s_1(t)$ は大域的な背景信号，$s_2(t)$ は雷雲からの放電，$s_3(t)$ は人工信号，$s_4(t)$ は外れ値の多い異常信号を想定している．混合行列 $\boldsymbol{A}$ は次のように定義した (図 6.3).

$$\boldsymbol{A} = \begin{bmatrix} 1.1098 & 0.0086 & 0.0000 & 0 \\ 1.0144 & 0.2795 & 0.0427 & 0 \\ 1.2989 & 0.1357 & 0.0097 & 0 \\ 1.0774 & 0.0000 & 0.0025 & 1 \\ 1.3034 & 0.0010 & 0.0000 & 0 \\ 1.0578 & 0.0000 & 0.0008 & 0 \\ 1.2241 & 0.0000 & 1.8119 & 0 \\ 1.0086 & 0.3309 & 0.0471 & 0 \\ 1.0101 & 0.0865 & 0.3639 & 0 \\ 1.2491 & 0.0255 & 0.0851 & 0 \\ 1.1325 & 1.9781 & 0.2574 & 0 \\ 1.3107 & 0.0003 & 0.0000 & 0 \end{bmatrix} \tag{6.14}$$

図の左奥へ延びる軸は源信号のインデックスを，右奥へ延びる軸は観測点のインデックスを，縦軸は混合行列 $\boldsymbol{A}$ の値 (対数スケール) を表している．また，大域信号 $s_1(t)$ を赤色で，異常信号 $s_4(t)$ を黄色で示している．

源信号を混合し，さらに絶対値ガウス雑音を加えることで生成した観測信号を図 6.4 に示す．図の軸は図 6.2 と同様である．どの観測信号にも 1 周期分の正弦波状の変化がほぼ同じ大きさでみられるが，これは大域信号 $s_1(t)$ の影響である．また，$x_4(t)$ には外れ値が多く見られるが，これは異常信号 $s_4(t)$ の影響である．

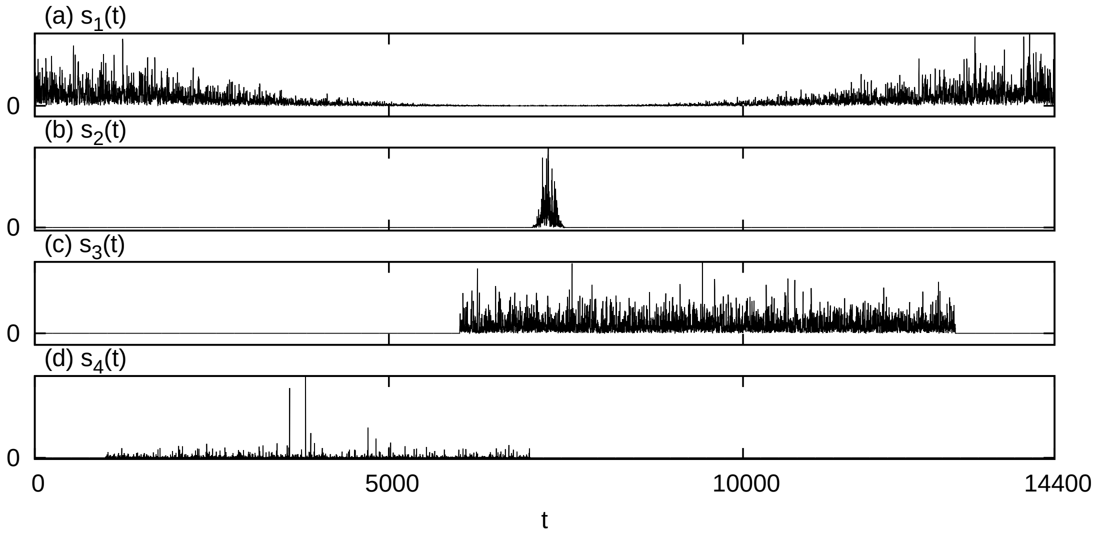

図 6.2: 生成した源信号

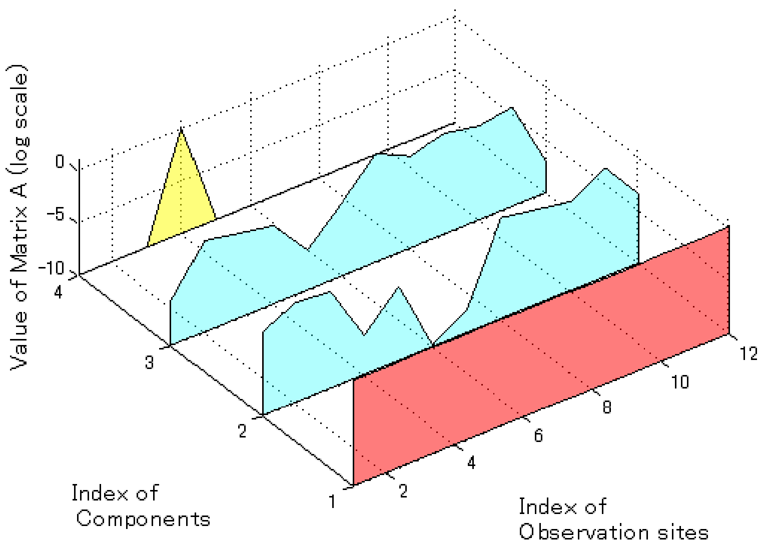

図 6.3: 生成した混合行列

図 6.5 および図 6.6 に，それぞれ ISRA および Quasi-L1 NMF によって推定した源信号を示す．図の軸は図 6.2 と同様である．どちらの手法でも源信号が良く推定できているように見える．しかし，推定された混合行列には大きな違いがある．図 6.7 および図 6.8 に，それぞれ ISRA および Quasi-L1 NMF によって推定した混合行列を示す．図の軸と色は図 6.3 と同様である．大域信号 $s_1(t)$ に対応する列 (赤色) において，ISRA によるものでは，$s_4(t)$ を多く含む No.4 の観測点に対応する値がほぼ 0 となっている．一方，Quasi-L1 NMF によるものでは他とほぼ同値であり，真の混合行列 (図 6.3) に良く似ている．

源信号の推定精度は，真の源信号との波形的な類似度だけでなく，混合行列も総合して評価すべきである．そこで，源信号 $s_j(t)$ に対応する推定値 $\hat{s}_k(t)$ に対し，次に定義する指標 C_j

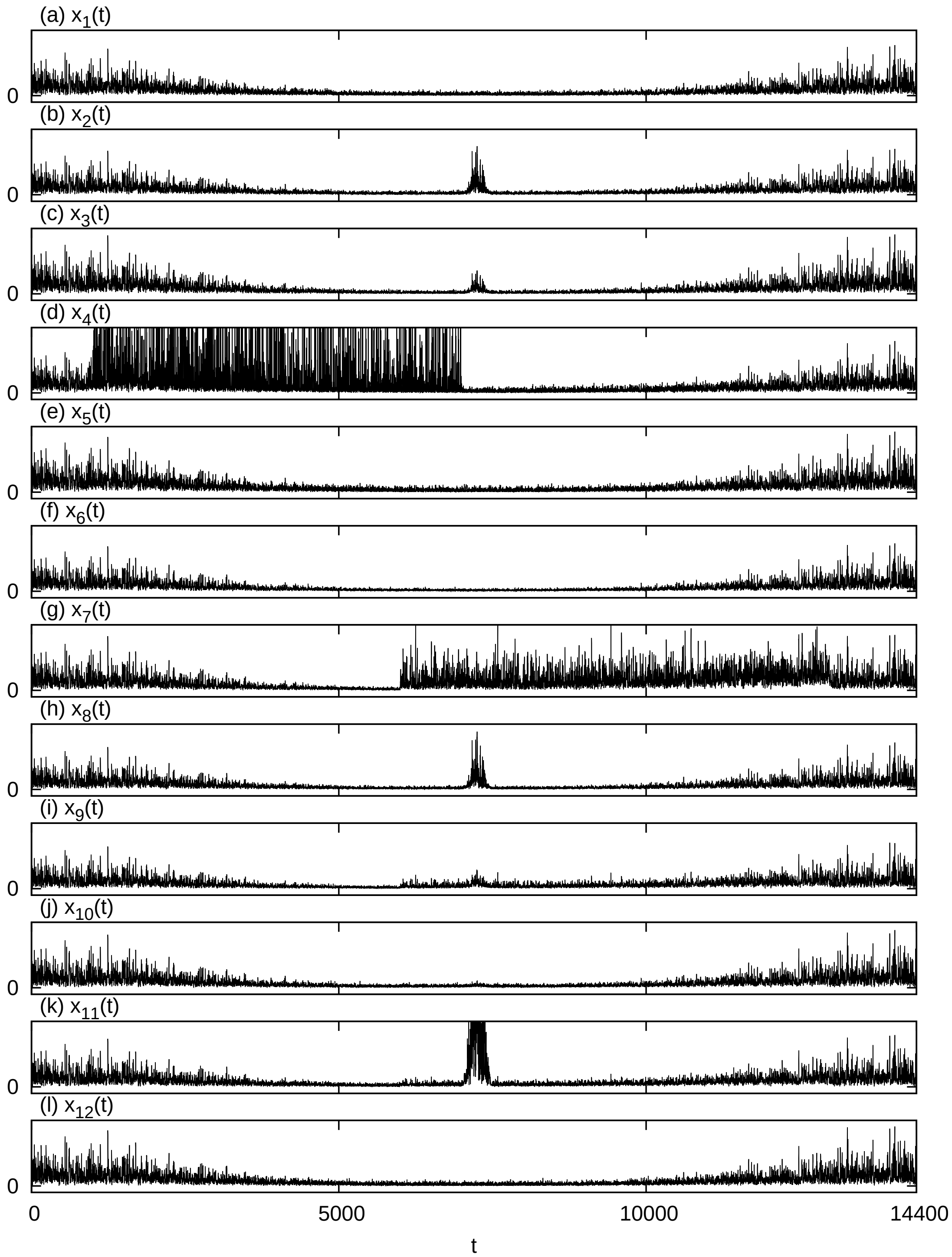

図 6.4: 生成した観測信号

を用いてその推定精度を評価する．

$$C_j \equiv 10\log_{10}\left\{\frac{1}{I}\sum_{i=1}^{I}\frac{V\{A_{ij}s_j(t)\}}{V\{A_{ij}s_k(t)-\hat{A}_{ik}\hat{s}_k(t)\}}\right\} \tag{6.15}$$

ここで，I は観測点数，$V\{\cdot\}$ は分散，$\hat{A}_{ik}$ は A_{ij} の推定値である．C_j は源信号 $s_j(t)$ の SN 比

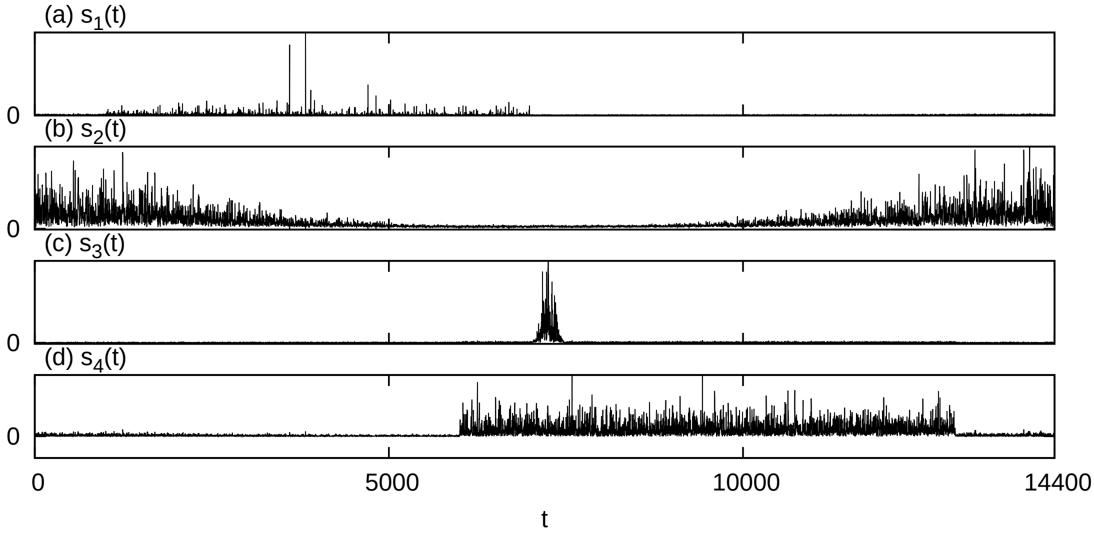

図 6.5: ISRA によって推定された源信号

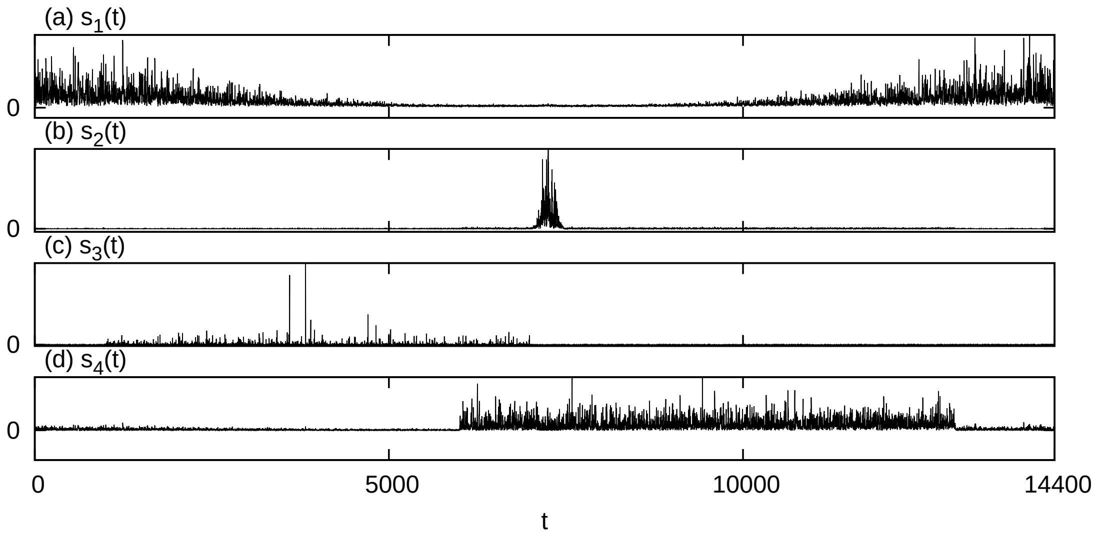

図 6.6: Quasi-L1 NMF によって推定された源信号

表 6.1: 各手法における源信号推定精度 C_j

手法 ＼ 源信号	$s_1(t)$	$s_2(t)$	$s_3(t)$	$s_4(t)$
ISRA	22.456	16.475	9.189	5.879
Quasi-L1 NMF	23.315	16.002	7.117	18.006

を観測点ごとに算出し，その平均を [dB] で表したものである．よって，C_j が大きな値であるほど推定精度が高いことを意味する．

各手法で推定した源信号の C_j を表 6.1 に示す．表より，源信号 $s_2(t)$ と $s_3(t)$ については Quasi-L1 NMF は ISRA に若干劣る結果となった．しかし，大域信号 $s_1(t)$ と外れ値の多い異常信号 $s_4(t)$ については，Quasi-L1 NMF の方が推定精度が高くなっている．特に $s_4(t)$ は，

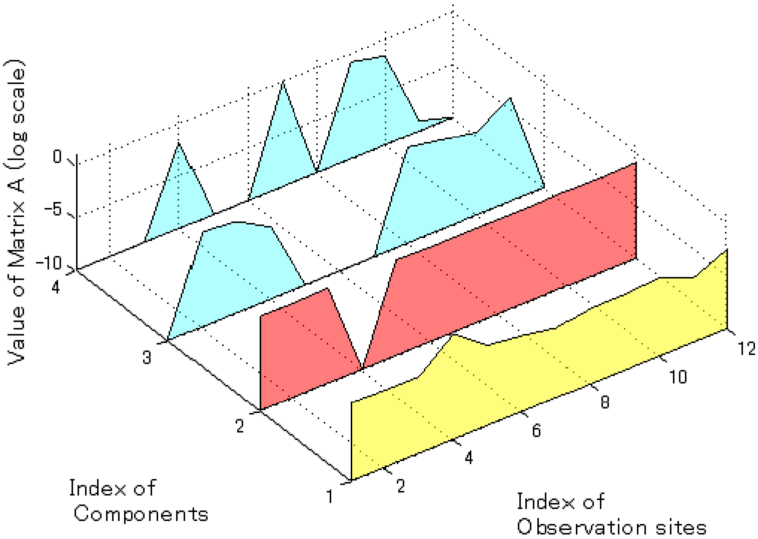

図 6.7: ISRA によって推定された混合行列

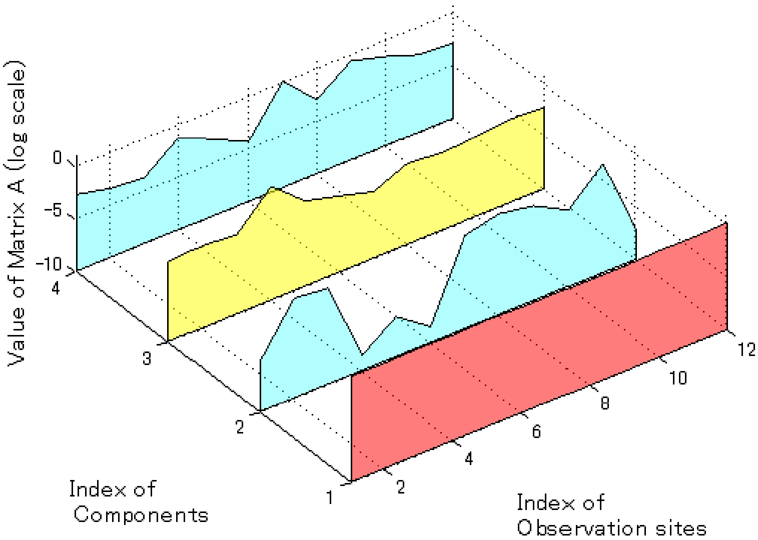

図 6.8: Quasi-L1 NMF によって推定された混合行列

10[dB] 以上改善していることから，Quasi-L1 NMF が外れ値を含む信号の解析に対し，頑強であることが分かる．

前章までに引き続き，大域信号除去の結果についても検証を行う．理想的な局所信号を図 6.9 に，ISRA および Quasi-L1 NMF によって推定された局所信号を，それぞれ図 6.10 および図 6.11 に示す．ISRA による局所信号では，(d) に示す $l_4(t)$ において，大域信号がほとんど除去されていない．これは，混合行列の推定において，大域信号の感度がほぼ 0 とされたためである．一方，Quasi-L1 NMF による局所信号では，(d) に示す $l_4(t)$ も含め，全ての局所信号が精度よく推定されている印象である．

推定された局所信号の精度を，SN 比および GIC に基づいて評価する (表 6.2)．表より，

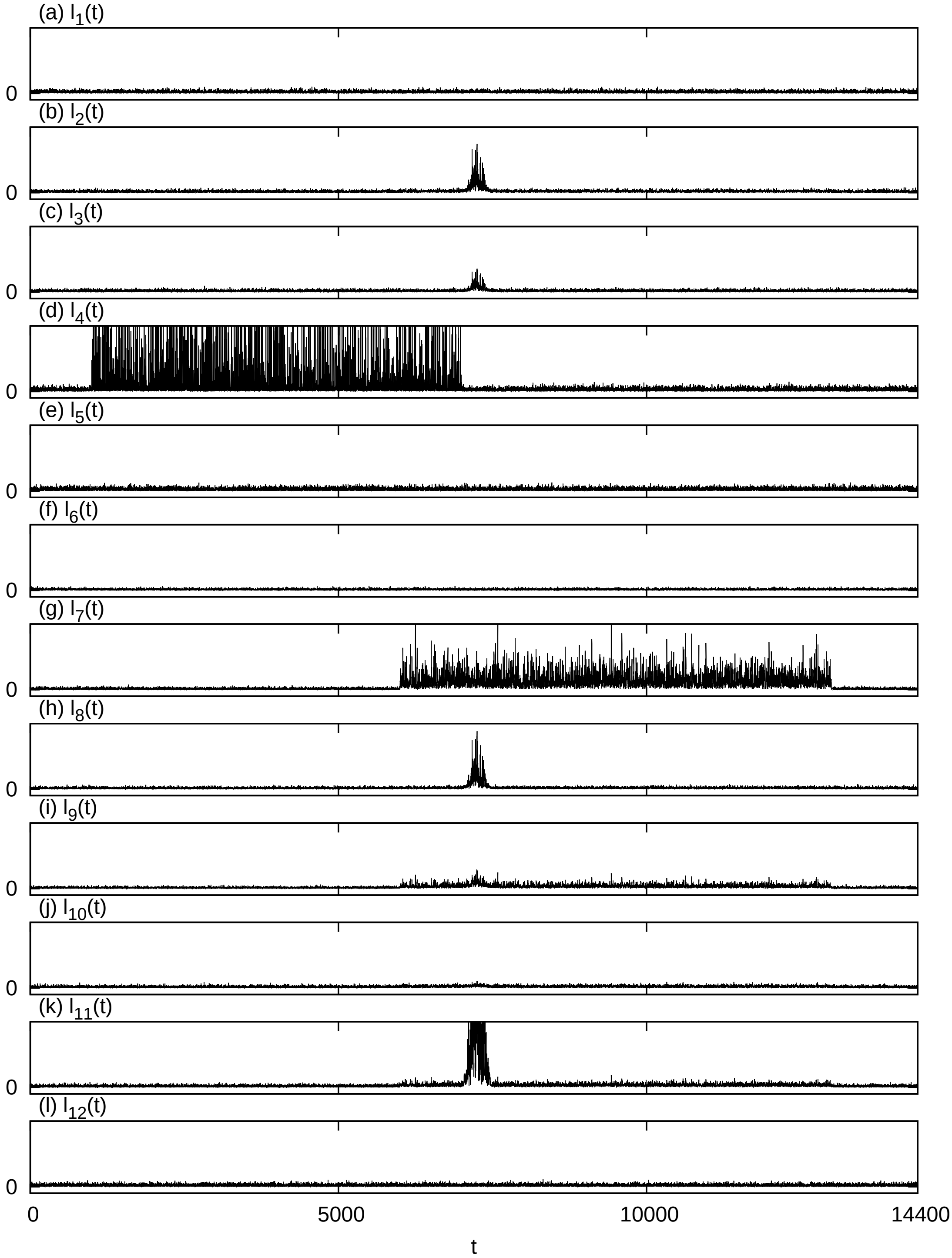

図 6.9: 理想的な局所信号

Quasi-L1 NMF による局所信号は，ISRA によるものと比較して，SN 比は約 10[dB] 改善しており，GIC も小さくなっている．これらの結果から，Quasi-L1 NMF は外れ値を多く含む観測信号からの大域信号除去において，ISRA よりも有効性が高いと言える．

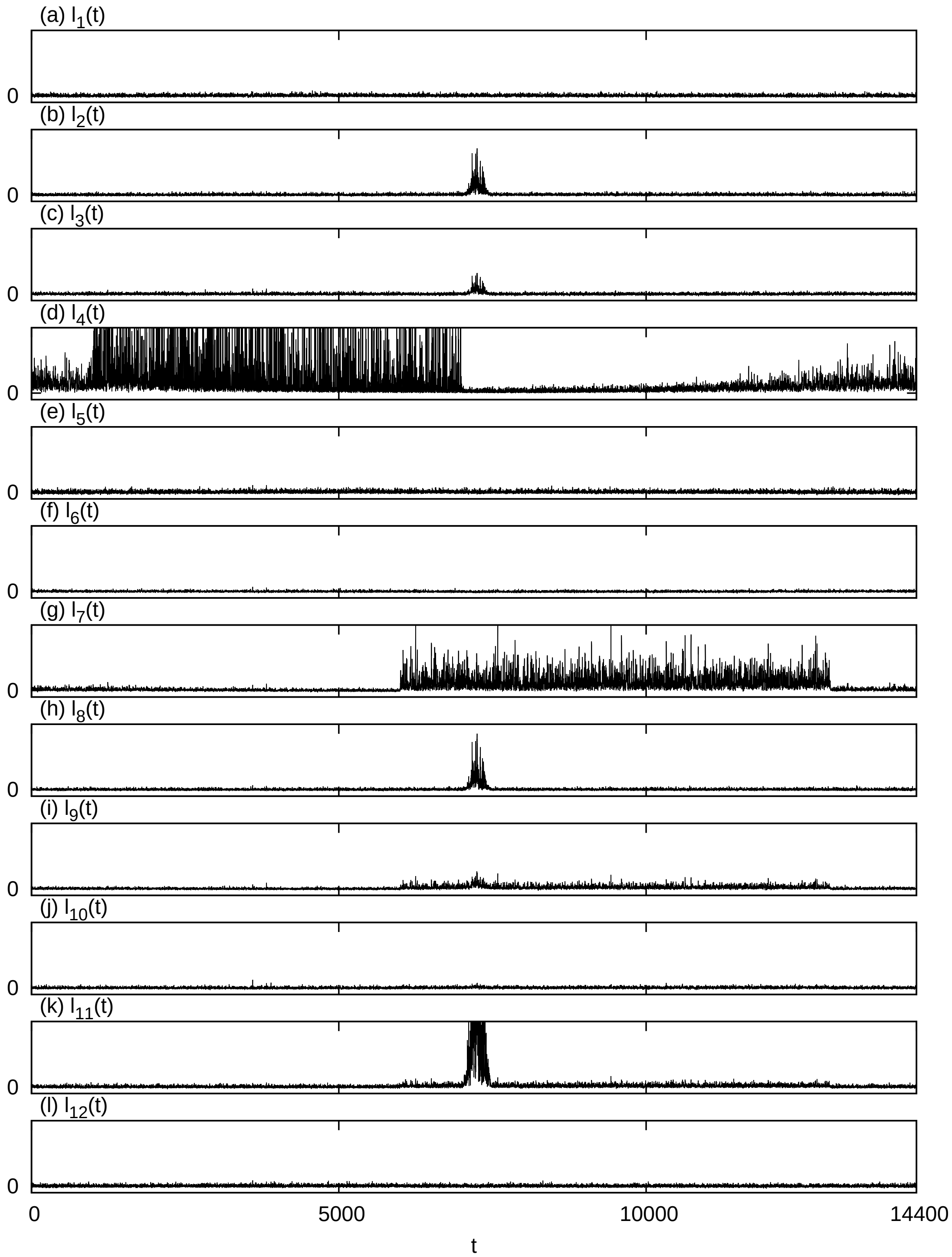

図 6.10: ISRA によって推定された局所信号

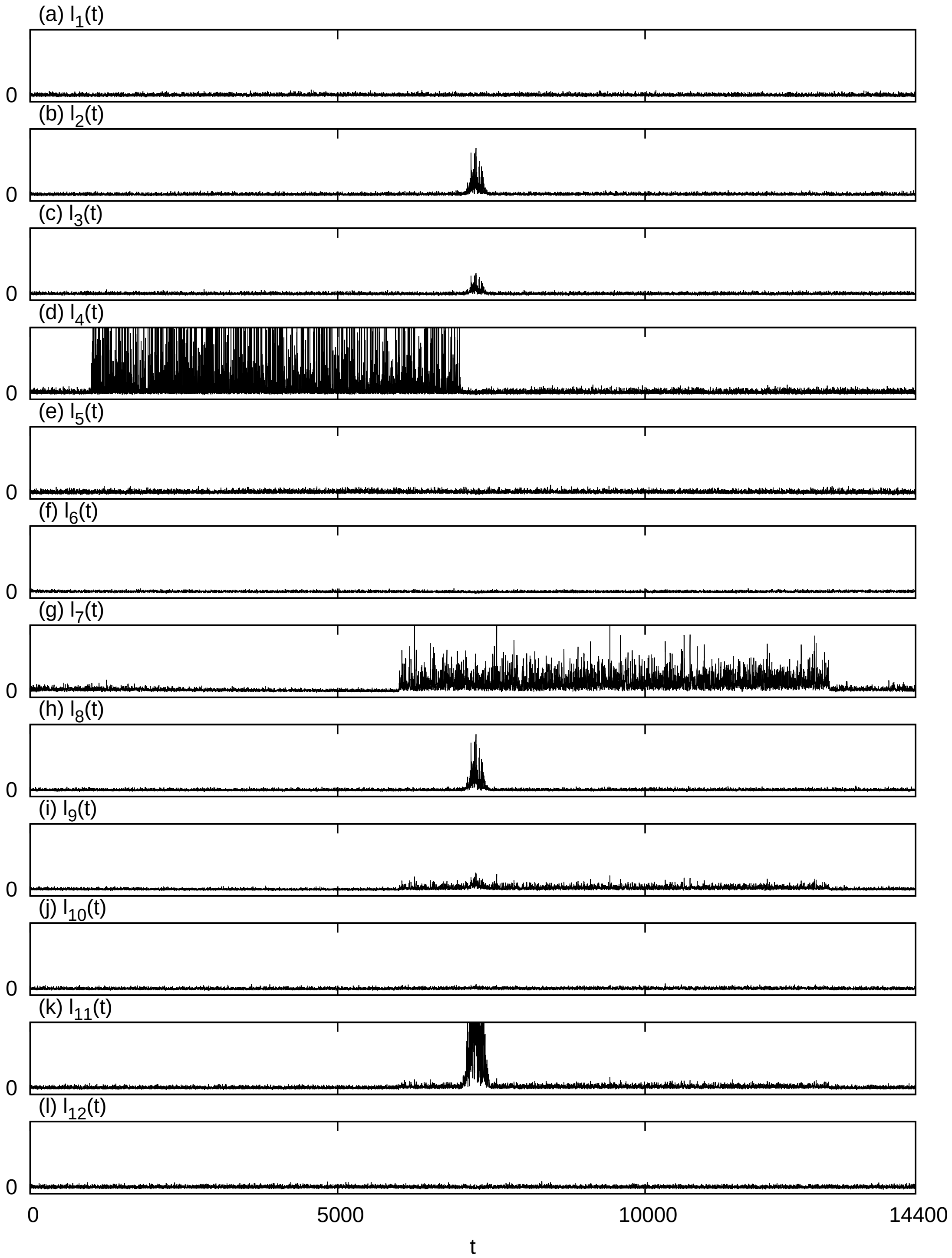

図 6.11: Quasi-L1 NMF によって推定された局所信号

表 6.2: 推定された局所信号の SN 比および GIC

データ	SNR[dB]	GIC
観測信号	6.997	0.742
推定された局所信号		
理想	—	0.061
平均的な観測信号	12.405	0.120
ISRA	21.484	0.071
Quasi-L1 NMF	30.960	0.068

6.4 ELF 帯環境電磁界観測データへの適用

図 6.12 は，2005 年 3 月 17 日における観測信号を一部抜粋したものである．図の (a) から (d) は，それぞれ (a) 秋田県男鹿，(b) 岐阜県萩原，(c) 大阪府茨木，(d) 長崎県雲仙において観測された信号の二乗値を表している．図の横軸は時系列 (日時) を，縦軸は ELF 帯環境電磁界のエネルギー [pT^2/Hz] を表している．観測信号には，共通した周期の長い変化が見られる．他のほとんどの観測点でも同様の変化が見られたことから，この変化は大域信号と呼ばれる，全国的に観測される背景信号によるものである．図中の (2) 長崎県雲仙市の信号には，午前中において多量のインパルス状の異常信号が見られる．この三日後の 3 月 20 日に福岡西方沖地震 (M7.0) が発生していることから，この異常信号が地震前兆を示している可能性が高い．

処理に用いるデータは，2005 年 3 月 17 日において，長崎県雲仙を含む 27 箇所の観測点にて観測された信号とした．提案した行列分解手法における出力信号数は任意であるが，本章においては出力信号数を 6 とした．

図 6.13 と図 6.14 に，ISRA によって推定された源信号と Quasi-L1 NMF によって推定された源信号を示す．図の縦軸は源信号の振幅を，横軸は時系列 (時刻) を表しており，図の (a) から (f) は，それぞれ $s_1(t)$ から $s_6(t)$ を表している．また，ISRA によって推定された混合行列と Quasi-L1 NMF によって推定された混合行列を，図 6.15 と図 6.16 に示す．

ISRA においては (d)$s_4(t)$ に，Quasi-L1 NMF においては (a)$s_1(t)$ に，観測信号に見られた共通の変化が良く推定されている．また，それぞれの混合行列における対応する列 (赤色) が，ほとんどの観測点でほぼ等しい値となってる．よって，大域信号だと断定できる．

ISRA においては (a)$s_2(t)$ に，Quasi-L1 NMF においては (f)$s_6(t)$ に，長崎県雲仙において観測された異常信号らしき成分が推定されている．また，それぞれの混合行列における対応する列 (黄色) が，No.23 の観測点 (長崎県雲仙) の値が突出している．しかし，ISRA の場合には大域信号らしき変化も含まれているように見える．そこで，波形の下部に着目して，これらを比較する．

図 6.17 に，推定された異常信号と岐阜県萩原の観測信号の，波形の下部を拡大したものを示す．(a) が ISRA によるもの (図 6.13(a)$s_2(t)$)，(b) が Quasi-L1 NMF によるもの (図 6.14(f)$s_6(t)$)，(c) が岐阜県萩原の観測信号 (図 6.12(b)) である．ISRA によって推定された異常信号の下部の包絡線は，長崎県雲仙から十分遠方である岐阜県萩原の観測信号の下部の包絡線とよく似ている．これは，ISRA によって推定された異常信号に大域信号の成分が含まれていることを，明確に示している．一方，Quasi-L1 NMF によって推定された異常信号の下部の包絡線は平坦である．よって，少なくとも大域信号の成分が含まれていないことを意

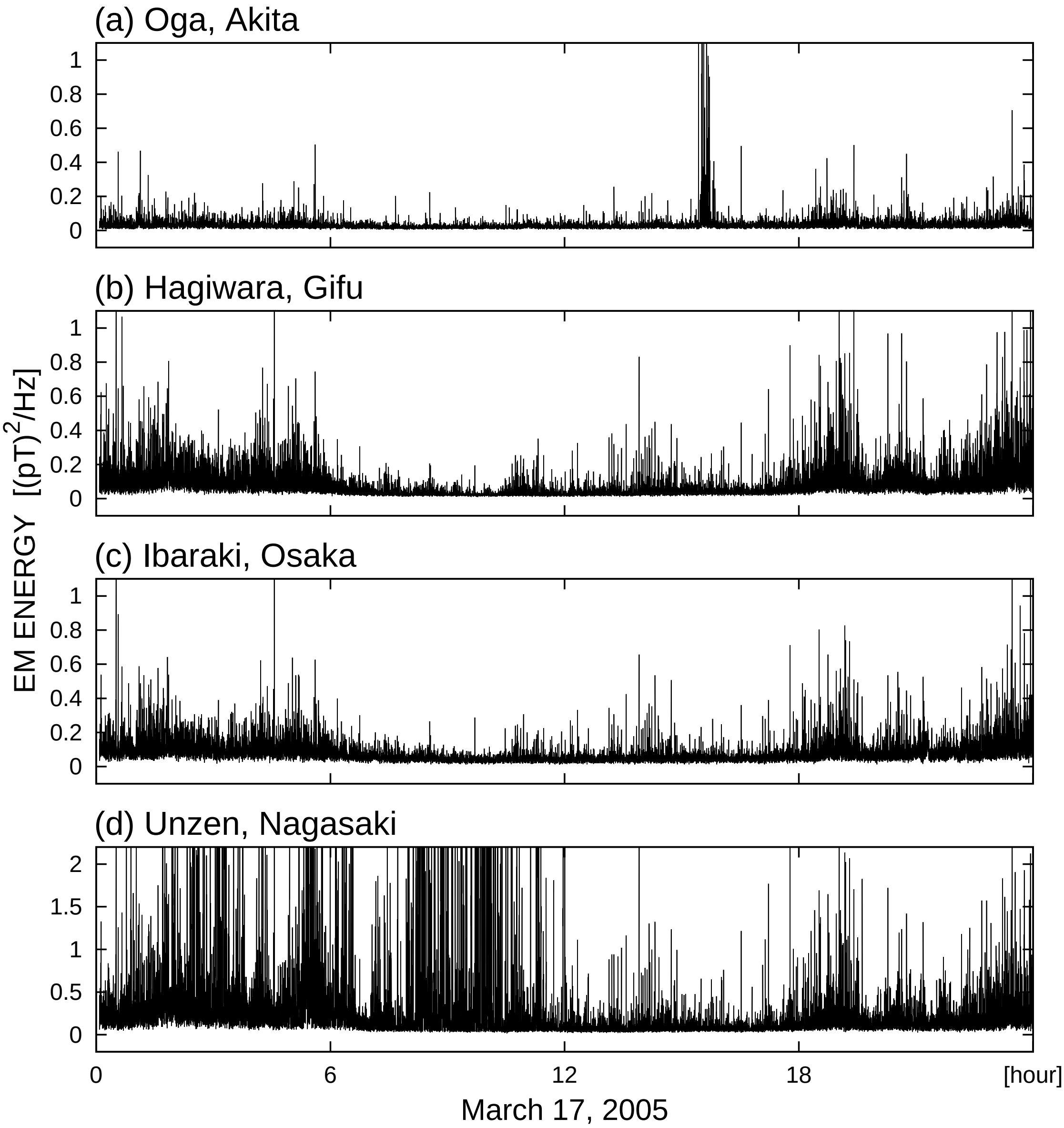

図 6.12: ELF 帯環境電磁波の観測信号 (2005 年 3 月 17 日，二乗値利用)

味している．また，間欠泉的に大きな値を取っている．これは，異常信号の放射が間欠泉的であることを意味している．地殻活動による電磁波放射がこのような性質をもつと仮定できる場合，Quasi-L1 NMF は，地殻活動によって放射された電磁波信号を，精度よく推定できていると言える．

前章までに引き続き，大域信号除去においての比較を行う．ISRA によって推定された局所信号を図 6.18 に示す．図の縦軸は源信号の振幅を，横軸は図 6.12 と同じく時系列 (時刻) を表している．大域信号が適切に除去された場合，観測信号に見られた共通の変化が見られなくなっているはずである．推定された局所信号を見ると，(a) から (c) の観測点に関して，共通した変化が見られなくなっている．しかし，外れ値の多い (d) 長崎県雲仙に関しては，観

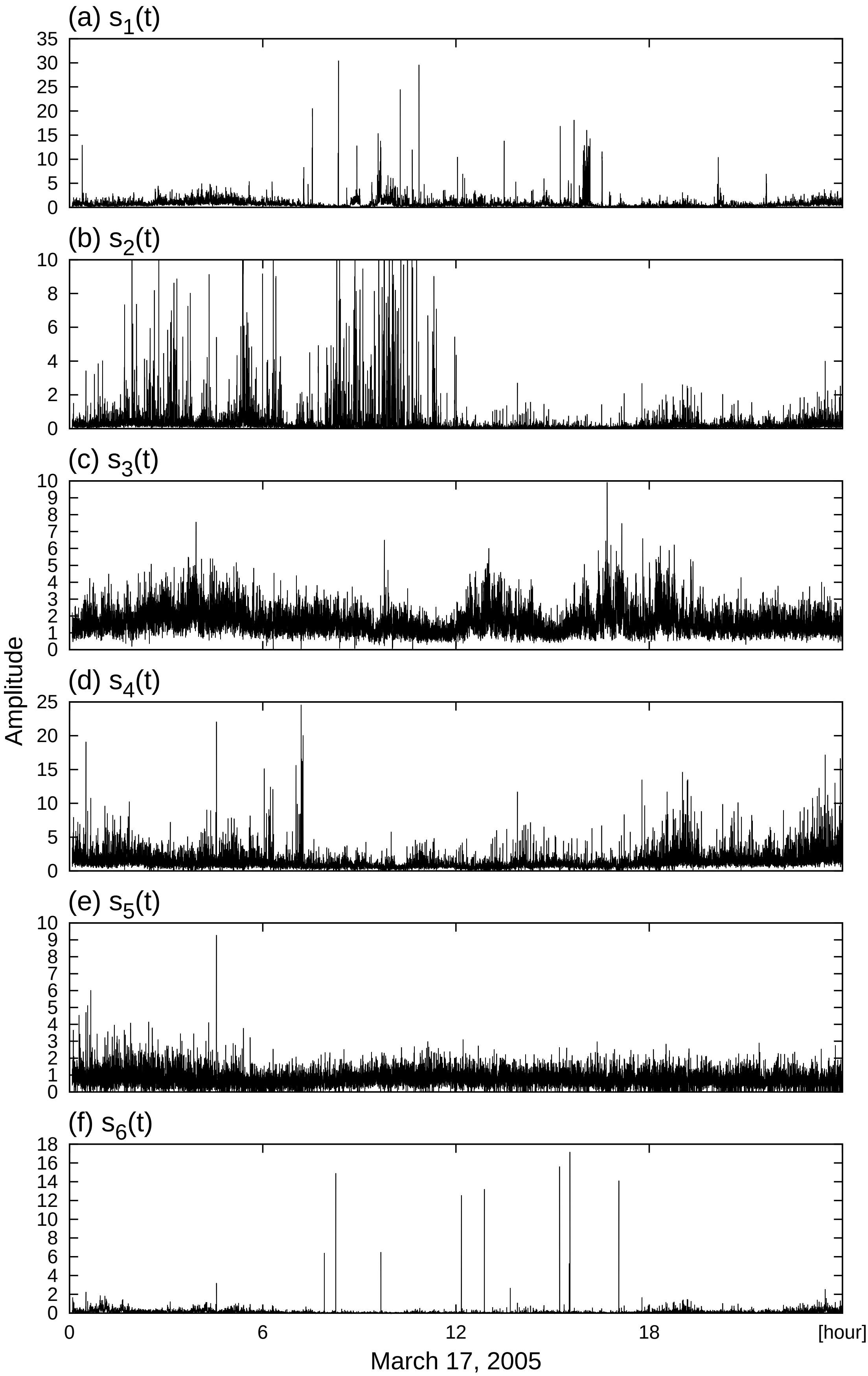

図 6.13: ISRA によって推定された源信号 (2005 年 3 月 17 日，二乗値利用)

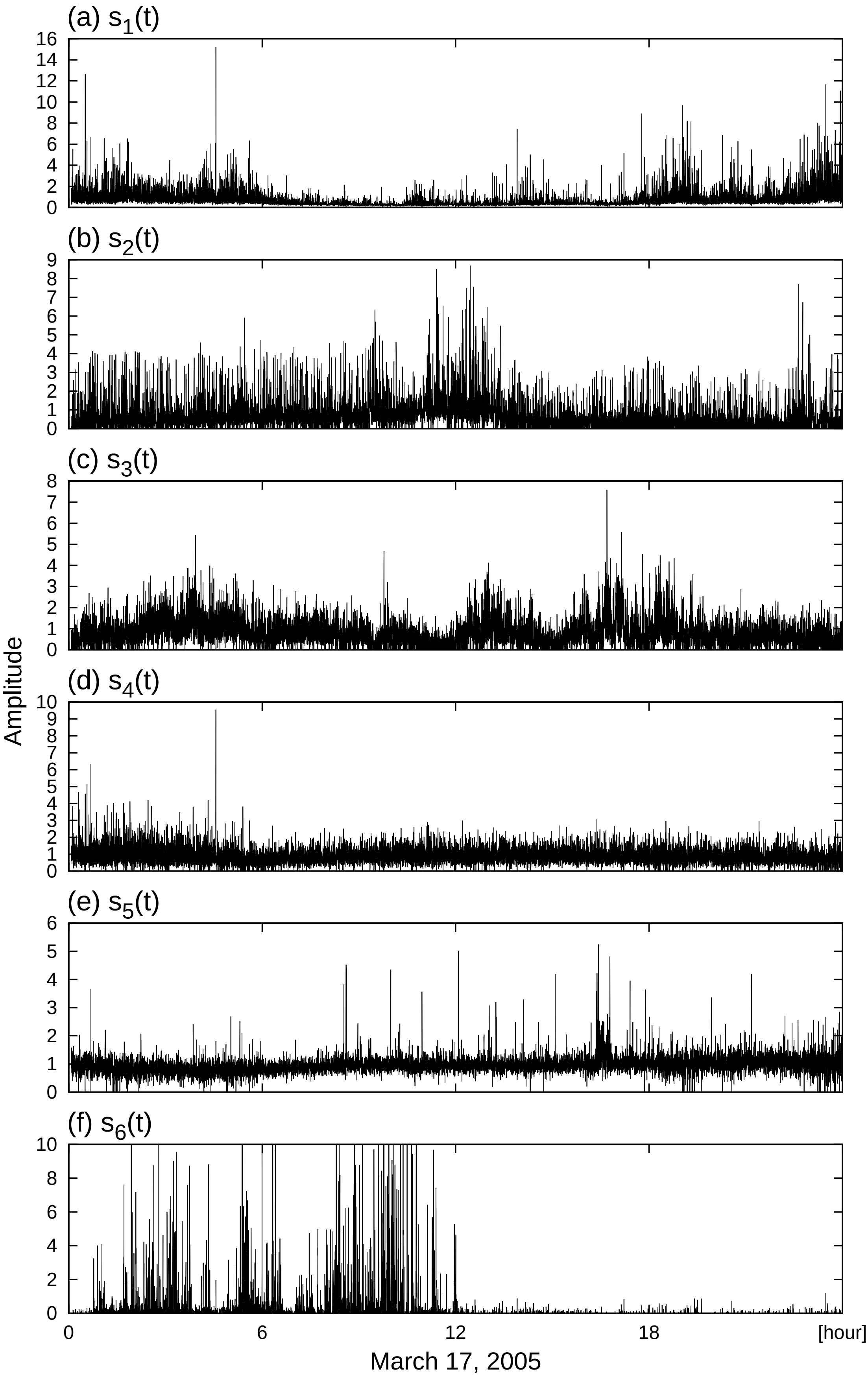

図 6.14: Quasi-L1 NMF によって推定された源信号 (2005 年 3 月 17 日，二乗値利用)

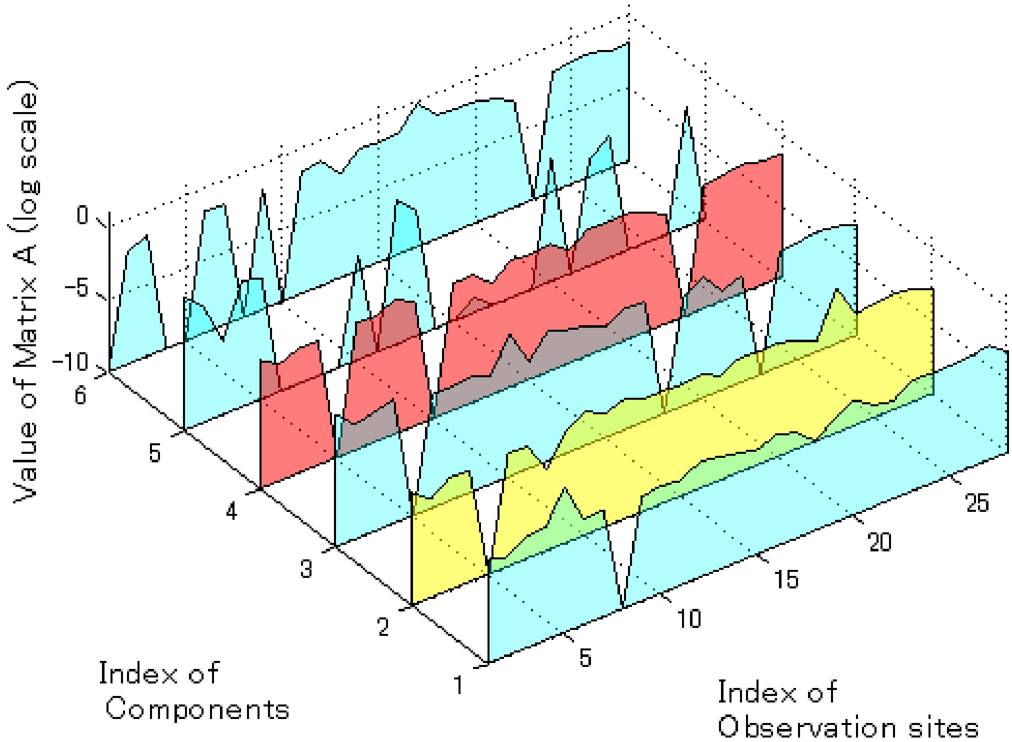

図 6.15: ISRA によって推定された混合行列 (2005 年 3 月 17 日，二乗値利用)

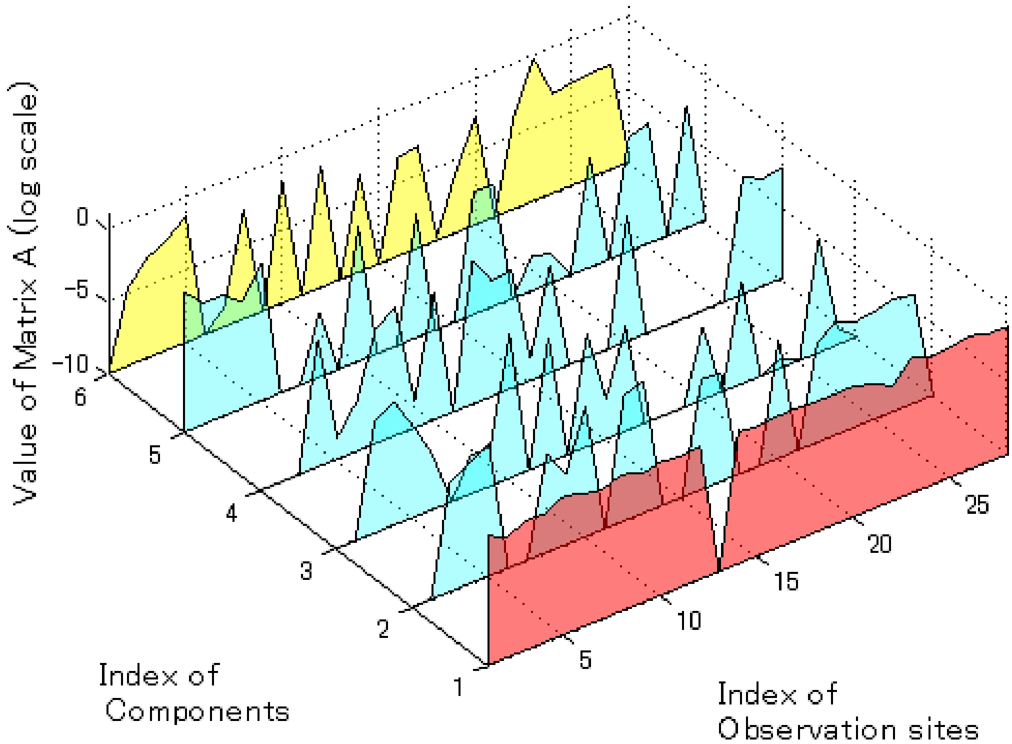

図 6.16: Quasi-L1 NMF によって推定された混合行列 (2005 年 3 月 17 日，二乗値利用)

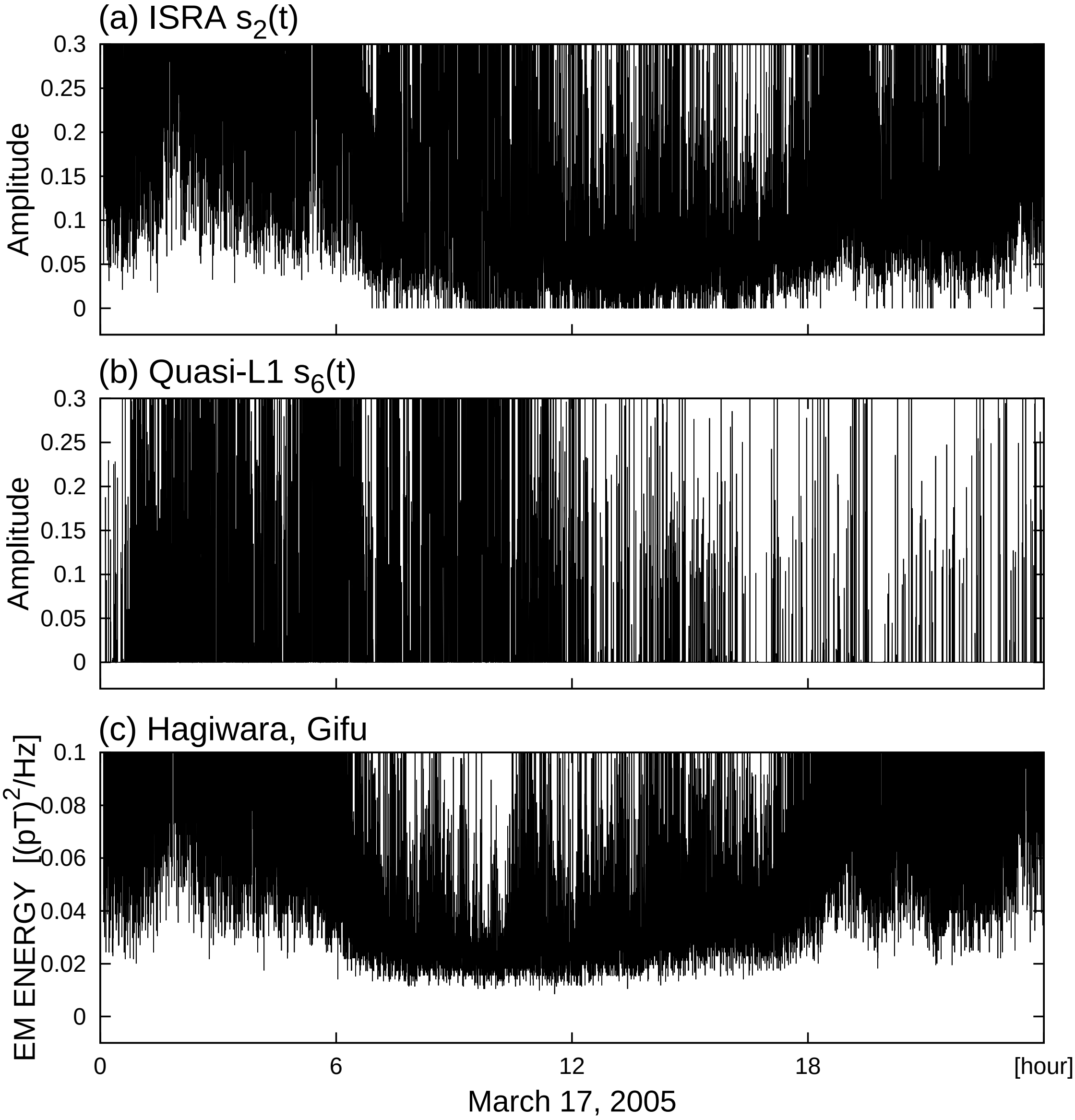

図 6.17: 推定された異常信号と岐阜県萩原の観測信号 (波形の下部を拡大)

表 6.3: 大域信号除去結果の評価 (2005 年 3 月 17 日，二乗値利用)

データ	GIC
観測信号	0.2447
推定された局所信号	
平均的な観測信号	0.1137
ISRA	0.0927
Quasi-L1 NMF	0.0763

測信号からの変化がほぼ無く，大域信号らしき成分が取り除かれていないと言える．これは，長崎県雲仙の観測信号に対する大域信号の感度の推定に失敗したためと考えられる．

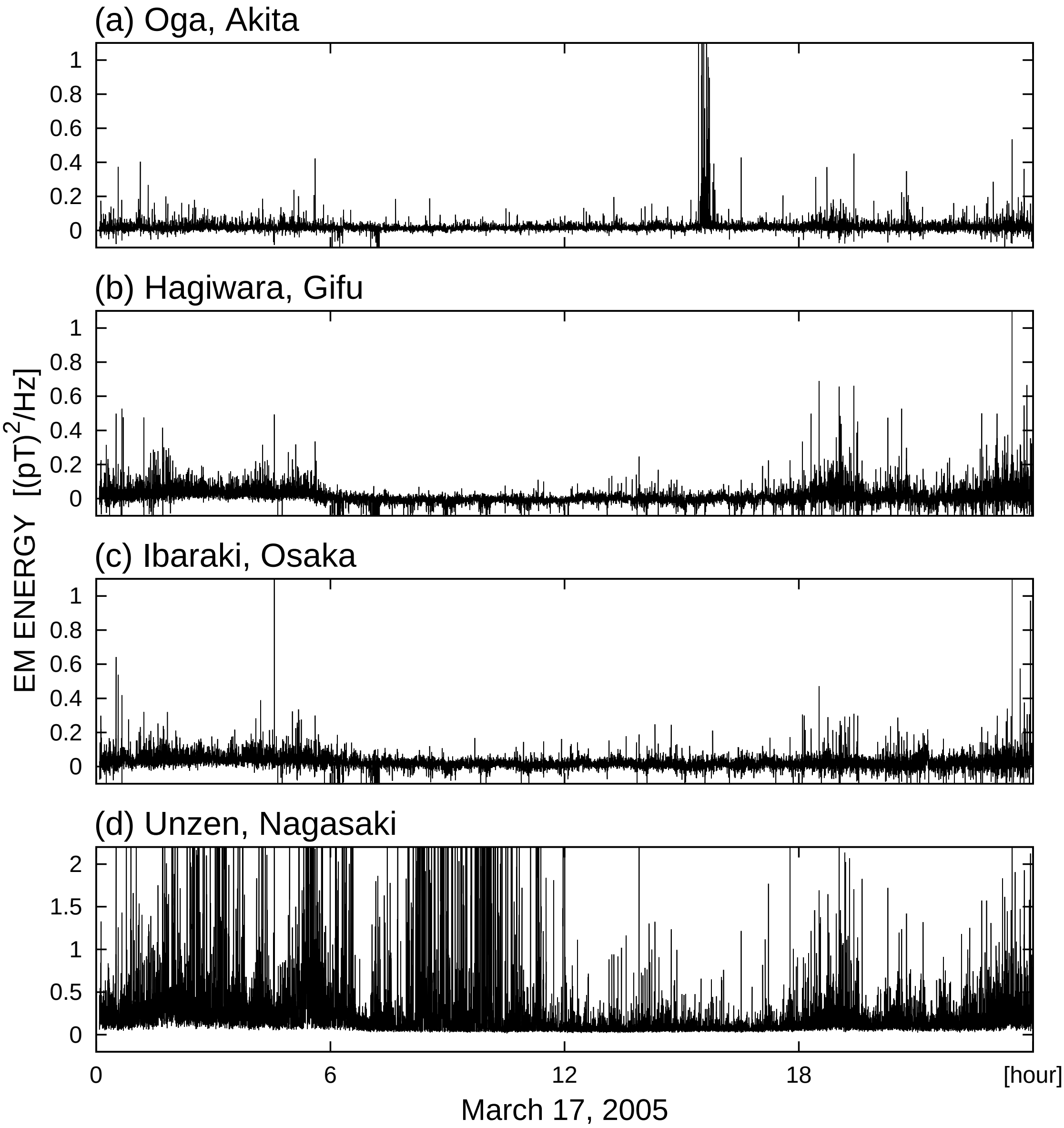

図 6.18: ISRA によって推定された局所信号 (2005 年 3 月 17 日，二乗値利用)

Quasi-L1 NMF によって推定された局所信号を図 6.19 に示す．推定された局所信号を見ると，(a) から (c) の観測点に関しては ISRA とほぼ同様であり，さらに，外れ値の多い (d) 長崎県雲仙に関しても，大域信号らしき成分が取り除かれている．また，ISRA によって推定された局所信号と比較すると，負の方向に大きな値を取る場合が減少している．

GIC を用いてそれぞれの信号を評価した (表 6.3)．表より，Quasi-L1 NMF の GIC は，ISRA を用いて推定した場合よりも小さい．このことから，Quasi-L1 NMF は精度良く大域信号が推定できていると言える．また，擬似信号を用いた実験を踏まえれば，そのほかの源信号も精度良く推定できていることが期待できる．

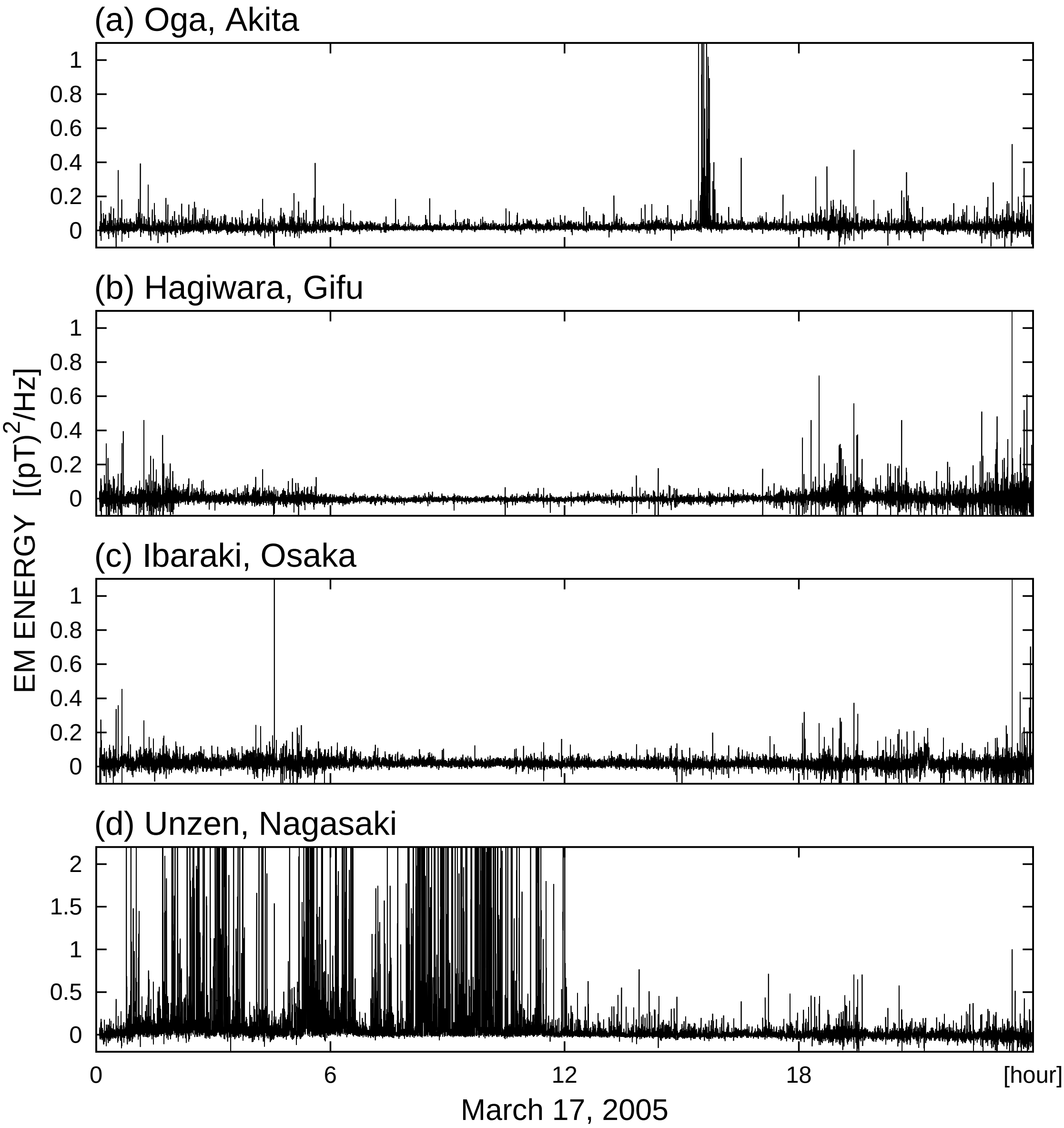

図 6.19: Quasi-L1 NMF によって推定された局所信号 (2005 年 3 月 17 日，二乗値利用)

6.5　まとめ

擬似 L1 ノルム最小化に基づいた非負行列因子分解手法を提案し，それを用いて環境電磁界観測データを模した擬似信号から源信号推定の実験を行った．また，実際の環境電磁界観測データに対し源信号の推定を行い，外れ値を多く含む異常信号の比較とともに，大域信号の推定精度を評価した．その結果，外れ値を多く含む異常信号の推定精度が大幅に向上すること，従来と同等もしくはより高い精度で大域信号を推定できることを示した．

今後の課題として，Quasi-L1 NMF の頑強性に関する定量的検証が挙げられる．また，環境電磁界観測データから Quasi-L1 NMF によって推定された，大域信号以外の信号について

追加の調査を行い，様々な知見と照らし合わせることも重要である．

加えて，Quasi-L1 NMF のアルゴリズムのさらなる改良も検討する必要がある．本論文において用いた ISRA および Quasi-L1 NMF では，特異値分解を用いて行列の初期化を行っている．これにより，大域信号などエネルギーの大きい源信号の推定精度が向上し，解も一意に定まっている．しかし，乱数を用いて行列を初期化した場合，源信号の推定精度が低下し，解も一意に定まらなくなる．つまり，現段階では推定結果が初期値に大きく依存することになる．これを解決すべく，新たな工夫を導入する必要がある．例えば，非負の信号における独立性の指標を模索し，アルゴリズムに組み込むことが考えられる．